Backward SDEs with constrained jumps and Quasi-Variational Inequalities

Idris Kharroubi, Jin Ma, Huyen Pham, Jianfeng Zhang

Backward SDEs with constrained jumps and Quasi-Variational Inequalities

Idris KHARROUBI[*] Jin MA[†] Huyên PHAM[‡] Jianfeng ZHANG[§]

May 30, 2008

Abstract

We consider a class of backward stochastic differential equations (BSDEs) driven by Brownian motion and Poisson random measure, and subject to constraints on the jump component. We prove the existence and uniqueness of the minimal solution for the BSDEs by using a penalization approach. Moreover, we show that under mild conditions the minimal solutions to these constrained BSDEs can be characterized as the unique viscosity solution of quasi-variational inequalities (QVIs), which leads to a probabilistic representation for solutions to QVIs. Such a representation in particular gives a new stochastic formula for value functions of a class of impulse control problems. As a direct consequence we obtain a numerical scheme for the solution of such QVIs via the simulation of the penalized BSDEs.

Key words : Backward stochastic differential equation, jump-diffusion process, jump constraints, penalization, quasi-variational inequalities, impulse control problems, viscosity solutions.

MSC Classification (2000): 60H10, 60H30, 35K85.

[*]Laboratoire de Probabilités et Modèles Aléatoires, CNRS, UMR 7599, Universités Paris 6-Paris 7, and CREST, kharroubi@ensae.fr

[†]Department of Mathematics, University of Southern California, Los Angeles, CA 90089, jinma@usc.edu. This author is supported in part by NSF grant#0505427.

[‡]Laboratoire de Probabilités et Modèles Aléatoires, CNRS, UMR 7599, Universités Paris 6-Paris 7, CREST, and Institut Universitaire de France, pham@math.jussieu.fr

[§]Department of Mathematics, University of Southern California, Los Angeles, CA 90089, jianfenz@usc.edu. This author is supported in part by NSF grant #0631366.

1 Introduction and summary

Consider a parabolic quasi-variational inequality (QVI for short) of the following form:

$$\min\Big[-\frac{\partial v}{\partial t} - \mathcal{L}v - f\ ,\ v - \mathcal{H}v\Big] \;=\; 0, \quad \text{on}\ \ [0,T)\times\mathbb{R}^d, \quad v(T,\cdot) \;=\; g \quad \text{on}\ \mathbb{R}^d, \tag{1.1}$$

where $\mathcal{L}$ is the second order local operator

$$\mathcal{L}v(t,x) \;=\; \langle b(x), D_x v(t,x)\rangle + \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x)D_x^2 v(t,x)) \tag{1.2}$$

and $\mathcal{H}$ is the nonlocal operator

$$\mathcal{H}v(t,x) \;=\; \sup_{e\in E}[v(t,x+\gamma(x,e)) + c(x,e)]. \tag{1.3}$$

In the above, $D_x v$ and $D_x^2 v$ are the partial gradient and the Hessian matrix of v with respect to its second variable x, respectively; $^{\intercal}$ stands for the transpose; $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathbb{R}^d$; $\mathbb{S}^d$ is the set of all symmetric $d\times d$ matrices; and E is some compact subset of $\mathbb{R}^q$.

It is well-known (see, e.g., [3]) that the QVI (1.1) is the dynamic programming equation associated to the impulse control problems whose value function is defined by:

$$v(t,x) \;=\; \sup_{\alpha=(\tau_i,\xi_i)_i} \mathbf{E}\Big[g(X_T^{t,x,\alpha}) + \int_t^T f(X_s^{t,x,\alpha})ds + \sum_{t<\tau_i\le T} c(X_{\tau_i^-}^{t,x,\alpha},\xi_i)\Big]. \tag{1.4}$$

More precisely, given a filtered probability space $(\Omega,\mathcal{F},\mathbf{P},\mathbb{F})$ where $\mathbb{F}=\{\mathcal{F}_t\}_t$, we define an impulse control α as a double sequence $(\tau_i,\xi_i)_i$ in which $\{\tau_i\}$ is an increasing sequence of $\mathbb{F}$-stopping times, and each ξ_i is an $\mathcal{F}_{\tau_i}$-measurable random variable taking values in E. For each impulse control $\alpha=(\tau_i,\xi_i)_i$, the controlled dynamics starting from x at time t, denoted by $X^{t,x,\alpha}$, is a càdlàg process satisfying the following SDE:

$$X_s^{t,x,\alpha} = x + \int_t^s b(X_u^{t,x,\alpha})du + \int_t^s \sigma(X_u^{t,x,\alpha})dW_u + \sum_{t<\tau_i\le s}\gamma(X_{\tau_i^-}^{t,x,\alpha},\xi_i), \tag{1.5}$$

where W is a d-dimensional $\mathbb{F}$-Brownian motion. In other words, the controlled process $X^{t,x,\alpha}$ evolves according to a diffusion process between two successive intervention times τ_i and τ_{i+1}, and at each decided intervention time τ_i, the process jumps with size $\Delta X_{\tau_i}^{t,x,\alpha} := X_{\tau_i}^{t,x,\alpha} - X_{\tau_i^-}^{t,x,\alpha} = \gamma(X_{\tau_i^-}^{t,x,\alpha},\xi_i)$.

We note that the impulse control problem (1.4) may be viewed as a sequence of optimal stopping problems combined with jumps in state due to impulse values. Moreover, the QVI (1.1) is the infinitesimal derivation of the dynamic programming principle, which means that at each time, the controller may decide either to do nothing and let the state process diffuse, or to make an intervention on the system via some impulse value. The former is characterized by the linear PDE in (1.1), while the latter is expressed by the obstacle (or reflected) part in (1.1). From the theoretical and numerical point of view, the main difficulty of the QVI (1.1) lies in that the obstacle contains the solution itself, and it is nonlocal (see

(1.3)) due to the jumps induced by the impulse control. These features make the classical approach of numerically solving such impulse control problems particular challenging.

An alternative method to attack the QVI (1.1) is to find the probabilistic representation of the solution using the Backward Stochastic Differential Equations (BSDEs), namely the so-called nonlinear Feynman-Kac formula. One can then hope to use such a representation to derive a direct numerical procedure for the solution of QVIs, whence the impulse control problems. The idea is the following. We consider a Poisson random measure $\mu(dt, de)$ on $\mathbb{R}_+ \times E$ associated to a marked point process $(T_i, \zeta_i)_i$. Assume that μ is independent of W and has intensity $\lambda(de)dt$, where λ is a finite measure on E. Consider a (uncontrolled) jump-diffusion process

$$X_s = X_0 + \int_0^s b(X_u)du + \int_0^s \sigma(X_u)dW_u + \sum_{T_i \le s} \gamma(X_{T_i^-}, \zeta_i). \tag{1.6}$$

Assume that v is a "smooth" solution to (1.1), and define $Y_t = v(t, X_t)$. Then, by Itô's formula we have

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T f(X_s)ds + K_T - K_t - \int_t^T \langle Z_s, dW_s \rangle \\ &\quad - \int_t^T \int_E (U_s(e) - c(X_{s^-}, e))\mu(ds, de), \end{aligned} \tag{1.7}$$

where $Z_t = \sigma^\intercal(X_{t^-})D_x v(t, X_{t^-})$, $U_t(e) = v(t, X_{t^-} + \gamma(X_{t^-}, e)) - v(t, X_{t^-}) + c(X_{t^-}, e)$, and $K_t = \int_0^t (-\dfrac{\partial v}{\partial t} - \mathcal{L}v - f)(s, X_s)ds$. Since v satisfies (1.1), we see that K is a continuous (hence predictable), nondecreasing process, and U satisfies the constraint:

$$-U_t(e) \geq 0, \tag{1.8}$$

The idea is then to view (1.7) and (1.8) as a BSDE with jump constraints, and we expect to retrieve $v(t, X_t)$ by solving the "minimal" solution (Y, Z, U, K) to this constrained BSDE.

We can also look at the BSDE above slightly differently. Let us denote $d\bar{K}_t = dK_t - \int_E U_s(e)\mu(dt, de)$, $t \geq 0$. Then $\bar{K}$ is still a nondecreasing process, and the equation (1.7) can now be rewritten as

$$Y_t = g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds, de) - \int_t^T \langle Z_s, dW_s \rangle + \bar{K}_T - \bar{K}_t. \tag{1.9}$$

We shall prove that $v(t, X_t)$ can also be retrieved by looking at the minimal solution $(Y, Z, \bar{K})$ to this BSDE. In fact, the following relation holds (assuming $t = 0$):

$$\begin{aligned} v(0, X_0) &= \inf\{y \in \mathbb{R} \ : \exists Z, \\ &\quad y + \int_0^T \langle Z_s, dW_s \rangle \geq g(X_T) + \int_0^T f(X_s)ds + \int_0^T \int_E c(X_{s^-}, e)\mu(ds, de)\}. \end{aligned} \tag{1.10}$$

We should mention that (1.10) also has a financial interpretation. That is, $v(0, x)$ is the minimal capital allowing to superhedge the payoff $\Pi_T(X) = g(X_T) + \int_0^T f(X_s)ds + \int_0^T c(X_{s^-}, e)\mu(ds, de)$ by trading only the asset W. Here, the market is obviously incomplete, since the jump part of the underlying asset X is not hedgeable.

Inspired by the above discussion, we now introduce the following general BSDE:

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds + K_T - K_t \qquad (1.11) \\ &\quad - \int_t^T \langle Z_s, dW_s \rangle - \int_t^T \int_E (U_s(e) - c(X_{s^-}, Y_{s^-}, Z_s, e))\mu(ds, de), \quad 0 \le t \le T, \end{aligned}$$

with constraints on the jump component in the form:

$$h(U_t(e)) \;\; \ge \;\; 0, \qquad \forall e \in E, \; 0 \le t \le T, \qquad (1.12)$$

where h is a given nonincreasing function. The solution to the BSDE is a quadruple (Y, Z, U, K) where, besides the usual the component (Y, Z, U), the fourth component K is a nondecreasing, càdlàg, adapted process, null at zero, which makes the constraint (1.12) possible. We note that without the constraint (1.12), the BSDE with $K = 0$ was studied by Tang and Li [21] and Barles, Buckdahn and Pardoux [2]. However, with the presence of the constraint, we may not have the uniqueness of the solution. We thus look only for the minimal solution (Y, Z, U, K), in the sense that for any other solution $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K})$ satisfying (1.11)-(1.12), it must hold that $Y \le \tilde{Y}$. Clearly, this BSDE is a generalized version of (1.7)-(1.8), where the functions f and c are independent of y and z, and $h(u) = -u$.

We can also consider the counterpart of (1.9), namely finding the minimal solution (Y, Z, K) of the BSDE:

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds + \int_t^T \int_E c(X_{s^-}, Y_{s^-}, Z_s, e)\mu(ds, de) \\ &\quad - \int_t^T \langle Z_s, dW_s \rangle + K_T - K_t, \qquad 0 \le t \le T. \qquad (1.13) \end{aligned}$$

It is then conceivable, as we shall prove, that this problem is a special case of (1.11)-(1.12) with $h(u) = -u$.

It is worth noting that if the generator f and the cost function c do not depend on y, z, which we refer to as the impulse control case, the existence of a minimal solution to the constrained BSDEs (1.7)-(1.8) may be directly obtained by supermartingale decomposition method in the spirit of El Karoui and Quenez [11] for the dual representation of the super-replication cost of $\Pi_T(X)$. In fact, the results could be extended easily to the case where f is linear in z, via a simple application of the Girsanov transformation. In our general case, however, we shall follow a penalization method, as was done in El Karoui et al. [10]. Namely, we construct a suitable sequence (Y^n, Z^n, U^n, K^n) of BSDEs with jumps, and prove that it converges to the minimal solution that we are looking for, by using a weak compactness argument of Peng [18].

Our next task of this paper is to relate the minimal solution to the BSDE with constrained jumps to the viscosity solutions to the following general QVI:

$$\min \Big[-\frac{\partial v}{\partial t} - \mathcal{L}v - f(\cdot, v, \sigma^{\intercal} D_x v) \; , \; h(\mathcal{H}v - v) \Big] \;\; = \;\; 0, \qquad (1.14)$$

where $\mathcal{H}$ is the nonlocal semilinear operator

$$\mathcal{H}v(t,x) \;\; = \;\; \sup_{e \in E} [v(t, x + \gamma(x,e)) + c(x, v(t,x), \sigma^{\intercal}(x) D_x v(t,x), e)].$$

Under suitable assumptions, we shall also prove the uniqueness of the viscosity solution, leading to a new probabilistic representation for this parabolic QVI.

We should point out that the BSDEs constraints have been studied by many authors. For example, El Karoui et al. [10] studied the reflected BSDEs, in which the component Y is forced to stay above a given obstacle; Cvitanic, Karatzas and Soner [8], and later Buckdahn and Hu [6] considered the case where the constraints are imposed on the component Z. Recently Peng [18] (see also [19]) studied the the general case where constraints are given on both Y and Z, which relates these constrained BSDEs to variational inequalities. The main feature of this work is to consider constraints on the jump component (U) of the solution, and to relate these jump-constrained BSDEs to quasi-variational inequalities. On the other hand, the classical approach in the theory and numerical approximation of impulse control problems and QVIs is to consider them as obstacle problems and iterated optimal stopping problems. However, our penalization procedure for jump-constrained BSDEs suggests a non-iterative approximation scheme for QVIs, which, to our best knowledge, is new.

The rest of the paper is organized as follows: In Section 2 we give a detailed formulation of BSDEs with constrained jumps, and show how it includes problem (1.13) as special case. Moreover, in the special case of impulse control, we directly construct and show the existence of a minimal solution. In Section 3 we develop the penalization approach for studying the existence of a minimal solution to our constrained BSDE for general f, c, and h. We show in Section 4 that the minimal solution to this constrained BSDE provides a probabilistic representation for the unique viscosity solution to a parabolic QVI. In Section 5 we discuss numerical issues for approximating QVIs by a penalization procedure. Finally, in Section 6 we provide some examples of sufficient conditions under which our general assumptions are satisfied.

2 BSDEs with constrained jumps

2.1 General formulation

Throughout this paper we assume that $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space on which are defined a d-dimensional standard Brownian motion $W = (W_t)_{t\geq 0}$, and a Poisson random measure μ on $\mathbb{R}_+ \times E$, where E is a compact set of $\mathbb{R}^q$, endowed with its Borel field $\mathcal{E}$. We assume that the Poisson random measure μ is independent of W, and has the intensity measure $\lambda(de)dt$ for some finite measure λ on $(E, \mathcal{E})$. We shall often assume that the support of λ is the whole space E, i.e.

$$\forall e \in E, \ \exists \mathcal{O} \ \text{ open neighborhood of } \ e, \ \lambda(\mathcal{O}) > 0.$$

We set $\tilde{\mu}(dt, de) = \mu(dt, de) - \lambda(de)dt$, the compensated measure associated to μ; and denote by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ the augmentation of the natural filtration generated by W and μ, and by $\mathcal{P}$ the σ-algebra of predictable subsets of $\Omega \times [0, T]$.

Given Lipschitz functions $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, and a measurable map $\gamma : \mathbb{R}^d \times E \to \mathbb{R}^d$, satisfying for some positive constants C and k_γ,

$$\sup_{e\in E} |\gamma(x, e)| \ \leq \ C, \quad \text{and} \quad \sup_{e\in E} |\gamma(x, e) - \gamma(x', e)| \ \leq \ k_\gamma |x - x'|, \quad x, x' \in \mathbb{R}^d,$$

we consider the forward SDE:

$$dX_s = b(X_s)ds + \sigma(X_s)dW_s + \int_E \gamma(X_{s^-}, e)\mu(ds, de). \tag{2.1}$$

Existence and uniqueness of (2.1) given an initial condition $X_0 \in \mathbb{R}^d$, is well-known under the above assumptions, and for any $0 \leq T < \infty$, we have the standard estimate

$$\mathbf{E}\Big[\sup_{0\leq t\leq T} |X_t|^2\Big] < \infty. \tag{2.2}$$

In what follows we fix a finite time duration $[0, T]$. Let us introduce some additional notations. We denote by

- $\mathcal{S}^\mathbf{2}$ the set of real-valued càdlàg adapted processes $Y = (Y_t)_{0\leq t\leq T}$ such that $\|Y\|_{\mathcal{S}^\mathbf{2}} := \Big(\mathbf{E}\Big[\sup_{0\leq t\leq T} |Y_t|^2\Big]\Big)^{\frac{1}{2}} < \infty$.
- $\mathbf{L^p(0, T)}$, $p \geq 1$, the set of real-valued processes $(\phi_t)_{0\leq t\leq T}$ such that $\mathbf{E}\Big[\int_0^T |\phi_t|^p dt\Big] < \infty$; and $\mathbf{L^p_{\mathbb{F}}(0, T)}$ is the subset of $\mathbf{L^p(0, T)}$ consisting of adapted processes.
- $\mathbf{L^p(W)}$, $p \geq 1$, the set of $\mathbb{R}^d$-valued $\mathcal{P}$-measurable processes $Z = (Z_t)_{0\leq t\leq T}$ such that $\|Z\|_{\mathbf{L^p(W)}} := \Big(\mathbf{E}\Big[\int_0^T |Z_t|^p dt\Big]\Big)^{\frac{1}{p}} < \infty$.
- $\mathbf{L^p}(\tilde{\mu})$, $p \geq 1$, the set of $\mathcal{P} \otimes \mathcal{E}$-measurable maps $U : \Omega \times [0, T] \times E \to \mathbb{R}$ such that $\|U\|_{\mathbf{L^p}(\tilde{\mu})} := \Big(\mathbf{E}[\int_0^T \int_E |U_t(e)|^p \lambda(de)dt\Big]\Big)^{\frac{1}{p}} < \infty$.
- $\mathbf{A^2}$ the closed subset of $\mathcal{S}^\mathbf{2}$ consisting of nondecreasing processes $K = (K_t)_{0\leq t\leq T}$ with $K_0 = 0$.

We are given four objects: (i) a terminal function, which is a measurable function $g : \mathbb{R}^d \mapsto \mathbb{R}$ satisfying a growth linear condition

$$\sup_{x\in\mathbb{R}^d} \frac{|g(x)|}{1 + |x|} < \infty, \tag{2.3}$$

(ii) a generator function f, which is a measurable function $f : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ satisfying a growth linear condition

$$\sup_{(x,y,z)\in\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d} \frac{|f(x, y, z)|}{1 + |x| + |y| + |z|} < \infty, \tag{2.4}$$

and an uniform Lipschitz condition on (y, z), i.e. there exists a constant k_f such that for all $x \in \mathbb{R}^d$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$,

$$|f(x, y, z) - f(x, y', z')| \leq k_f(|y - y'| + |z - z'|), \tag{2.5}$$

(iii) a cost function, which is a measurable function $c : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times E \to \mathbb{R}$ satisfying a growth linear condition

$$\sup_{(x,y,z,e)\in\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times E} \frac{|c(x, y, z, e)|}{1 + |x| + |y| + |z|} < \infty, \tag{2.6}$$

and an uniform Lipschitz condition on (y,z), i.e. there exists a constant k_c such that for all $x \in \mathbb{R}^d$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$, $e \in E$,

$$|c(x,y,z,e) - c(x,y',z',e)| \leq k_c(|y-y'| + |z-z'|), \tag{2.7}$$

(iv) a constraint function, which is a measurable map $h : \mathbb{R} \times E \to \mathbb{R}$ s.t for all $e \in E$,

$$u \longmapsto h(u,e) \quad \text{is nonincreasing,} \tag{2.8}$$

and satisfying a Lipschitz condition on u i.e. there exists a constant k_h such that for all $u, u' \in \mathbb{R}$, $e \in E$,

$$|h(u,e) - h(u',e)| \leq k_h|u-u'|. \tag{2.9}$$

Let us now introduce our BSDE with constrained jumps: find a quadruple (Y,Z,U,K) $\in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ satisfying

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T f(X_s,Y_s,Z_s)ds + K_T - K_t \qquad (2.10)\\ &\quad - \int_t^T \langle Z_s, dW_s\rangle - \int_t^T \int_E (U_s(e) - c(X_{s^-},Y_{s^-},Z_s,e))\mu(ds,de), \quad 0 \leq t \leq T, \ a.s. \end{aligned}$$

with

$$h(U_t(e),e) \geq 0, \qquad d\mathbf{P} \otimes dt \otimes \lambda(de) \ \ a.e. \tag{2.11}$$

and such that for any other quadruple $(\tilde{Y},\tilde{Z},\tilde{U},\tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ satisfying (2.10)-(2.11), we have

$$Y_t \leq \tilde{Y}_t, \quad 0 \leq t \leq T, \ a.s.$$

We say that Y is the minimal solution to (2.10)-(2.11). In the formulation of Peng, one may sometimes say that Y is the smallest supersolution to (2.10)-(2.11). We shall also say that (Y,Z,U,K) is a minimal solution to (2.10)-(2.11), and we discuss later the uniqueness of such quadruple.

Remark 2.1 Since we are originally motivated by probabilistic representation of QVI's, we put the BSDE with constrained jumps in a Markovian framework. But all the results of Section 3 about the existence and approximation of a minimal solution hold true in a general non Markovian framework with the following standard modifications : the terminal condition $g(X_T)$ is replaced by a square integrable random variable $\xi \in \mathbf{L^2}(\boldsymbol{\Omega}, \mathcal{F}_\mathbf{T})$, the generator is a map f from $\Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d$ into $\mathbb{R}$, satisfying a uniform Lipschitz condition in (y,z), and $f(\cdot,y,z) \in \mathbf{L}^\mathbf{2}_\mathbb{F}(\mathbf{0},\mathbf{T})$ for all $(y,z) \in \mathbb{R} \times \mathbb{R}^d$, and the cost coefficient is a map c from $\Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times E$ into $\mathbb{R}$, satisfying a uniform Lipschitz condition in (y,z), and $c(\cdot,y,z,e) \in \mathbf{L}^\mathbf{2}_\mathbb{F}(\mathbf{0},\mathbf{T})$ for all $(y,z,e) \in \mathbb{R} \times \mathbb{R}^d \times E$.

Remark 2.2 Without the h-constraint condition (2.11) on jumps, we have existence and uniqueness of a solution (Y, Z, U, K) with $K = 0$ to (2.10), from results on BSDE with jumps in [21] and [2]. Here, under (2.11) on jumps, it is not possible in general to have equality in (2.10) with $K = 0$, and as usual in the BSDE literature with constraint, we consider a nondecreasing process K to have more freedom. The problem is then to find a minimal solution to this constrained BSDE, and the nondecreasing condition (2.8) on h is crucial for stating comparison principles needed in the penalization approach. The primary example of constraint function is $h(u, e) = -u$, i.e. nonpositive jumps constraint, which is actually equivalent to consider minimal solution to BSDE (1.13) as showed later.

2.2 The case of nonpositive jump constraint

Let us recall the BSDE defined in the introduction: find a triplet $(Y, Z, K) \in \mathcal{S}^{\mathbf{2}} \times \mathbf{L}^{\mathbf{2}}(\mathbf{W}) \times \mathbf{A}^{\mathbf{2}}$ such that

$$\begin{aligned} Y_t \;&=\; g(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds + K_T - K_t \\ &\quad - \int_t^T \langle Z_s, dW_s\rangle + \int_t^T \int_E c(X_{s^-}, Y_{s^-}, Z_s, e)\mu(ds, de), \quad 0 \le t \le T, \ a.s. \end{aligned} \tag{2.12}$$

such that for any other triplet $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^{\mathbf{2}} \times \mathbf{L}^{\mathbf{2}}(\mathbf{W}) \times \mathbf{A}^{\mathbf{2}}$ satisfying (2.12), it holds that

$$Y_t \;\le\; \tilde{Y}_t, \quad 0 \le t \le T, \ a.s.$$

We will call such Y (and, by a slight abuse of notation, (Y, Z, K)) the *minimal solution* to (2.12). We claim that this problem is actually equivalent to problem (2.10)-(2.11) in the case $h(u, e) = -u$, corresponding to nonpositive jump constraint condition:

$$U_t(e) \;\le\; 0, \qquad d\mathbf{P} \otimes dt \otimes \lambda(de) \ \ a.e. \tag{2.13}$$

Indeed, let (Y, Z, U, K) be any solution of (2.10) and (2.13). Define a process $\bar{K}$ by $d\bar{K}_t = dK_t - \int_E U_s(e)\mu(dt, de)$, $0 \le t \le T$, then $\bar{K}$ is nondecreasing, and the triplet $(Y, Z, \bar{K})$ satisfies (2.12). It follows that the minimal solution to (2.12) is smaller than the minimal solution to (2.10) and (2.13). We shall see in the next section, by using comparison principles and penalization approach, that equality holds, i.e.

$$\text{minimal solution } Y \text{ to (2.12)} \;=\; \text{minimal solution } Y \text{ to (2.10), (2.13).}$$

We shall illustrate this result by considering a special: when the functions f and c do not depend on y, z (i.e., the impulse control case). In this case, one can obtain directly the existence of a minimal solution to (2.10)-(2.13) and (2.12) by duality methods involving the following set of probability measures. Let $\mathcal{V}$ be the set of $\mathcal{P} \otimes \mathcal{E}$-measurable essentially bounded processes valued in $(0, \infty)$, and given $\nu \in \mathcal{V}$, consider the probability measure $\mathbf{P}^\nu$ equivalent to $\mathbf{P}$ on $(\Omega, \mathcal{F}_T)$ with Radon-Nikodym density :

$$\frac{d\mathbf{P}^\nu}{d\mathbf{P}} \;=\; \mathcal{E}_T\Big(\int_0^\cdot \int_E (\nu_t(e) - 1)\tilde{\mu}(dt, de)\Big), \tag{2.14}$$

where $\mathcal{E}_t(.)$ is the Doléans-Dade exponential. Notice that the Brownian motion W remains a Brownian motion under $\mathbf{P}^\nu$, which can then be interpreted as an equivalent martingale measure for the "asset" price process W. The effect of the probability measure $\mathbf{P}^\nu$, by Girsanov's theorem, is to change the compensator $\lambda(de)dt$ of μ under $\mathbf{P}$ to $\nu_t(e)\lambda(de)dt$ under $\mathbf{P}^\nu$.

In order to ensure that the problem is well-defined, we need to assume :

(H1) There exists a triple $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$ satisfying (2.12).

This assumption is standard and natural in the literature on BSDE with constraints, and means equivalently here (when f and c do not depend on y, z) that one can find some constant $\tilde{y} \in \mathbb{R}$, and $\tilde{Z} \in \mathbf{L^2(W)}$ such that

$$\tilde{y} + \int_0^T \langle \tilde{Z}_s, dW_s \rangle \;\; \geq \;\; g(X_T) + \int_0^T f(X_s)ds + \int_0^T \int_E c(X_{s^-}, e)\mu(ds, de) \quad a.s.$$

This equivalency can be proved by same arguments as in [8]. Notice that Assumption **(H1)** may be not satisfied as shown in Remark 3.1, in which case the problem (2.12) is ill-posed.

Theorem 2.1 *Suppose that f and c do not depend on y, z, and* **(H1)** *holds. Then, there exists a unique minimal solution $(Y, Z, K, U) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$, with K predictable, to* (2.10)-(2.13). *Moreover, $(Y, Z, \bar{K})$ is the unique minimal solution to* (2.12) *with $\bar{K}_t = K_t - \int_0^t \int_E U_s(e)\mu(ds, de)$, and Y has the explicit functional representation :*

$$Y_t \;\; = \;\; \operatorname*{ess\,sup}_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds, de) \Big| \mathcal{F}_t \Big], \quad 0 \leq t \leq T.$$

Proof. First, observe that for any $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ (resp. $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$) satisfying (2.10)-(2.13) (resp. (2.12)), the process

$$\tilde{Q}_t \;\; := \;\; \tilde{Y}_t + \int_0^t f(X_s)ds + \int_0^t \int_E c(X_{s^-}, e)\mu(ds, de), \quad 0 \leq t \leq T,$$

is a $\mathbf{P}^\nu$-supermartingale, for all $\nu \in \mathcal{V}$, where the probability measure $\mathbf{P}^\nu$ was defined in (2.14). Indeed, from (2.10)-(2.13) (resp. (2.12)), we have

$$\tilde{Q}_t \;\; = \;\; \tilde{Q}_0 + \int_0^t \langle \tilde{Z}_s, dW_s \rangle - \bar{K}_t, \quad \text{with} \quad \bar{K}_t = \tilde{K}_t - \int_0^t U_s(e)\mu(ds, de),$$

$$(\text{ resp.} \quad \tilde{Q}_t \;\; = \;\; \tilde{Q}_0 + \int_0^t \langle \tilde{Z}_s, dW_s \rangle - \tilde{K}_t), \quad 0 \leq t \leq T.$$

Now, by Girsanov's theorem, W remains a Brownian motion under $\mathbf{P}^\nu$, while from the boundedness of $\nu \in \mathcal{V}$, the density $d\mathbf{P}^\nu/d\mathbf{P}$ lies in $L^2(\mathbf{P})$. Hence, from Cauchy-Schwarz inequality, the condition $\tilde{Z} \in \mathbf{L^2(W)}$, and Burkholder-Davis-Gundy inequality, we get the $\mathbf{P}^\nu$-martingale property of the stochastic integral $\int \langle \tilde{Z}, dW \rangle$, and so the $\mathbf{P}^\nu$-supermartingale property of $\tilde{Q}$ since $\bar{K}$ (resp. $\tilde{K}$) is nondecreasing. This implies

$$\tilde{Y}_t \;\; \geq \;\; \mathbf{E}^\nu \Big[\tilde{Y}_T + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds, de) \Big| \mathcal{F}_t \Big],$$

and thereby, from the arbitrariness of $\mathbf{P}^\nu$, $\nu \in \mathcal{V}$, and since $\tilde{Y}_T = g(X_T)$,

$$Y_t := \operatorname*{ess\,sup}_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s) ds + \int_t^T \int_E c(X_{s^-}, e) \mu(ds, de) \Big| \mathcal{F}_t \Big] \quad \leq \quad \tilde{Y}_t. \quad (2.15)$$

To show the converse, let us consider the process Y defined in (2.15). By standard arguments as in [11], the process Y can be considered in its càd-làg modification, and we also notice that $Y \in \mathcal{S}^\mathbf{2}$. Indeed, by observing that the choice of $\nu = 1$ corresponds to the probability $\mathbf{P}^\nu = \mathbf{P}$, we have $\hat{Y} \leq Y \leq \tilde{Y}$, where $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$ is a solution to (2.12), and

$$\hat{Y}_t \quad = \quad \mathbf{E} \Big[g(X_T) + \int_t^T f(X_s) ds + \int_t^T \int_E c(X_{s^-}, e) \mu(ds, de) \Big| \mathcal{F}_t \Big].$$

Thus, since $\hat{Y}$ lies in $\mathcal{S}^\mathbf{2}$ from the linear growth conditions on g, f, and c, and the estimate (2.2), we deduce that $Y \in \mathcal{S}^\mathbf{2}$. Now, by similar dynamic programming arguments as in [11], we see that the process

$$Q_t \quad = \quad Y_t + \int_0^t f(X_s) ds + \int_0^t \int_E c(X_{s^-}, e) \mu(ds, de), \quad 0 \leq t \leq T, \quad (2.16)$$

lies in $\mathcal{S}^\mathbf{2}$, and is a $\mathbf{P}^\nu$-supermartingale, for all $\nu \in \mathcal{V}$. Then, from the Doob-Meyer decomposition of Q under each $\mathbf{P}^\nu$, $\nu \in \mathcal{V}$, we obtain :

$$Q_t \quad = \quad Y_0 + M^\nu - K^\nu, \quad (2.17)$$

where M^ν is a $\mathbf{P}^\nu$-martingale, $M^\nu = 0$, and K^ν is a $\mathbf{P}^\nu$ nondecreasing predictable càd-làg process with $K_0^\nu = 0$. Recalling that W is a $\mathbf{P}^\nu$-Brownian motion, and since $\tilde{\mu}^\nu(ds, de)$:= $\mu(ds, de) - \nu_s(e)\lambda(de)ds$ is the compensated measure of μ under $\mathbf{P}^\nu$, the martingale representation theorem for each M^ν, $\nu \in \mathcal{V}$ gives the existence of predictable processes Z^ν and U^ν such that

$$Q_t \quad = \quad Y_0 + \int_0^t \langle Z_s^\nu, dW_s \rangle + \int_0^t \int_E U_s^\nu(e) \tilde{\mu}^\nu(ds, de) - K_t^\nu, \quad 0 \leq t \leq T. \quad (2.18)$$

By comparing the decomposition (2.18) under $\mathbf{P}^\nu$ and $\mathbf{P}$ corresponding to $\nu = 1$, and identifying the martingale parts and the predictable finite variation parts, we obtain that $Z^\nu = Z^1 =: Z$, $U^\nu = U^1 =: U$ for all $\nu \in \mathcal{V}$, and

$$K_t^\nu \quad = \quad K_t^1 - \int_0^t \int_E U_s(e)(\nu_s(e) - 1)\lambda(de) ds, \quad 0 \leq t \leq T. \quad (2.19)$$

Now, by writing the relation (2.18) with $\nu = \varepsilon > 0$, substituting the definition of Q in (2.16), and since $Y_T = g(X_T)$, we obtain :

$$\begin{aligned} Y_t \quad = \quad & g(X_T) + \int_t^T f(X_s) ds - \int_t^T \langle Z_s, dW_s \rangle - \int_t^T \int_E (U_s(e) - c(X_{s^-}, e)) \mu(ds, de) \\ & + \int_t^T \int_E U_s(e) \varepsilon \lambda(de) ds + K_T^\varepsilon - K_t^\varepsilon, \quad 0 \leq t \leq T. \end{aligned} \quad (2.20)$$

From (2.19), the process K^ε has a limit as ε goes to zero, which is equal to $K^0 = K^1 + \int_0^\cdot \int_E U_s(e)\lambda(de)ds$, and inherits from K^ε, the nondecreasing path and predictability properties. Moreover, since $Q \in \mathcal{S}^\mathbf{2}$, in the decomposition (2.17) of Q under $\mathbf{P} = \mathbf{P}^\nu$ for $\nu = 1$, the process M^1 lies in $\mathcal{S}^\mathbf{2}$ and $K^1 \in \mathbf{A^2}$. This implies that $Z \in \mathbf{L^2(W)}$, $U \in \mathbf{L^2}(\tilde{\mu})$, and also that $K^0 \in \mathbf{A^2}$. By sending ε to zero into (2.20), we obtain that $(Y, Z, U, K^0) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ is a solution to (2.10). Let us finally check that U satisfies the constraint :

$$U_t(e) \quad \leq \quad 0, \quad d\mathbf{P} \otimes dt \otimes \lambda(de). \tag{2.21}$$

We argue by contradiction by assuming that the set $F = \{(\omega, t, e) \in \Omega \times [0,T] \times E \ : U_t(e) > 0\}$ has a strictly positive measure for $d\mathbf{P} \times dt \times \lambda(de)$. For any $k > 0$, consider the process $\nu_k = 1_{F^c} + (k+1)1_F$, which lies in $\mathcal{V}$. From (2.19), we have

$$\mathbf{E}[K_T^{\nu_k}] \quad = \quad \mathbf{E}[K_T^1] - k\mathbf{E}\Big[\int_0^T \int_E 1_F U_t(e)\lambda(de)dt\Big] \ < \ 0,$$

for k large enough. This contradicts the fact that $K_T^{\nu_k} \geq 0$, and so (2.21) is satisfied. Therefore (Y, Z, U, K^0) is a solution to (2.10)-(2.13), and it is a minimal solution from (2.15). Y is unique by definition. The uniqueness of Z follows by identifying the Brownian parts and the finite variation parts, and the uniqueness of (U, K^0) is obtained by identifying the predictable parts by recalling that the jumps of μ are inacessible. By denoting $\bar{K}^0 = K^0 - \int_0^t \int_E U_s(e)\mu(ds, de)$, which lies in $\mathbf{A^2}$, we see that $(Y, Z, \bar{K}^0)$ is a solution to (2.12), and it is minimal by (2.15). Uniqueness follows by identifying the Brownian parts and the finite variation parts. □

Remark 2.3 In Section 4, we shall relate rigorously the constrained BSDEs (2.10)-(2.11) to QVIs. In particular, the minimal solution Y_t to (2.10)-(2.13) or (2.12) is $Y_t = v(t, X_t)$ where v is the value function of the impulse control problem (1.4). Together with the functional representation of Y in Theorem 2.1, we then have the following relation at time $t = 0$:

$$v(0, X_0) \quad = \quad \sup_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T) + \int_0^T f(X_s)ds + \int_0^T \int_E c(X_{s^-}, e)\mu(ds, de)\Big]. \tag{2.22}$$

We then recover a recent result obtained by Bouchard [4], who related impulse controls to stochastic target problems in the case of a finite set E. We may also interpret this result as follows. Recall that the effect of the probability measure $\mathbf{P}^\nu$ is to change the compensator $\lambda(de)dt$ of μ under $\mathbf{P}$ to $\nu_t(e)\lambda(de)dt$ under $\mathbf{P}^\nu$. Hence, by taking the supremum over all $\mathbf{P}^\nu$, we formally expect to retrieve in distribution law all the dynamics of the controlled process in (1.5) when varying the impulse controls α, which is confirmed by the equality (2.22).

Finally, we mention that the above duality and martingale methods may be extended when the generator function f is linear in z by using Girsanov's transformation. Our main purpose is now to study the general case of h-constraints on jumps, and nonlinear functions f and c depending on y, z.

3 Existence and approximation by penalization

In this section, we prove the existence of a minimal solution to (2.10)-(2.11), based on approximation via penalization. For each $n \in \mathbb{N}$, we introduce the penalized BSDE with jumps

$$\begin{aligned} Y_t^n &= g(X_T) + \int_t^T f(X_s, Y_s^n, Z_s^n) ds + n \int_t^T \int_E h^-(U_s^n(e), e) \lambda(de) ds \qquad (3.1) \\ &\quad - \int_t^T \langle Z_s^n, dW_s \rangle - \int_t^T \int_E (U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)) \mu(ds, de), \quad 0 \leq t \leq T, \end{aligned}$$

where $h^-(u, e) = \max(-h(u, e), 0)$ is the negative part of the function h. Under the Lipschitz conditions on the coefficients f, c and h, we know from the theory of BS-DEs with jumps, see [21] and [2], that there exists a unique solution $(Y^n, Z^n, U^n) \in \mathcal{S}^2 \times \mathbf{L}^2(\mathbf{W}) \times \mathbf{L}^2(\tilde{\mu})$ to (3.1). We define for each $n \in \mathbb{N}$,

$$K_t^n = n \int_0^t \int_E h^-(U_s^n(e), e) \lambda(de) ds, \quad 0 \leq t \leq T,$$

which is a nondecreasing process in $\mathbf{A}^2$. The rest of this section is devoted to the convergence of the sequence $(Y^n, Z^n, U^n, K^n)_n$ to the minimal solution we are interested in.

3.1 Comparison results

We first state that the sequence $(Y^n)_n$ is nondecreasing. This follows from a comparison theorem for BSDEs with jumps whose generator is of the form $\tilde{f}(x, y, z, u) = f(x, y, z) + \int_E \tilde{h}(u(e), e) \lambda(de)$ for some nondecreasing function $\tilde{h}$, which covers our situation from the nonincreasing condition on the constraint function h.

Lemma 3.1 *The sequence $(Y^n)_n$ is nondecreasing, i.e. for all $n \in \mathbb{N}$, $Y_t^n \leq Y_t^{n+1}$, $0 \leq t \leq T$, a.s.*

Proof. Define the sequence $(V^n)_n$ of $\mathcal{P} \otimes \mathcal{E}$-measurable processes by

$$\begin{aligned} V_t^n(e) &= U_t^n(e) - c(X_{t^-}, Y_{t^-}^n, Z_t^n, e), \quad (t, e) \in (0, T] \times E \quad \text{and} \\ V_0^n(e) &= U_0^n(e) - c(X_0, Y_0^n, Z_0^n, e), \quad e \in E, \end{aligned}$$

From (3.1) and recalling that X and Y are càd-làg, we see that (Y^n, Z^n, V^n) is the unique solution in $\mathcal{S}^2 \times \mathbf{L}^2(\mathbf{W}) \times \mathbf{L}^2(\tilde{\mu})$ of the BSDE with jumps :

$$Y_t^n = g(X_T) + \int_t^T F_n(X_s, Y_s^n, Z_s^n, V_s^n) ds - \int_t^T \langle Z_s^n, dW_s \rangle - \int_t^T \int_E V_s^n(e) \tilde{\mu}(ds, de),$$

with $F_n(x, y, z, v) = f(x, y, z) + \int_E (v(e) + nh^-(v(e) + c(x, y, z, e), e)) \lambda(de)$. Since h^- is nondecreasing, we have

$$\begin{aligned} F_n(t, x, y, z, v) - F_n(t, x, y, z, v') &= \int_E \big\{ (v(e) - v'(e)) + n[h^-(v(e) + c(x, y, z, e), e) \\ &\qquad - h^-(v'(e) + c(x, y, z, e), e)] \big\} \lambda(de) \\ &\leq \int_E \big\{ (1 + \mathbf{1}_{\{v(e) \geq v'(e)\}} n k_h)(v(e) - v'(e)) \big\} \lambda(de). \end{aligned}$$

Moreover, since $F_{n+1} \geq F_n$, we can apply the comparison theorem 2.5 of [20], and obtain that $Y_t^n \leq Y_t^{n+1}$, $0 \leq t \leq T$, a.s. □

The next result shows that the sequence $(Y^n)_n$ is upper-bounded by any solution to the constrained BSDE. Arguments in the proof involve suitable change of probability measures $\mathbf{P}^\nu$, $\nu \in \mathcal{V}$, introduced in (2.14).

Lemma 3.2 *For any quadruple* $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ *satisfying* (2.10)-(2.11)*, and for all* $n \in \mathbb{N}$*, we have*

$$Y_t^n \quad \leq \quad \tilde{Y}_t, \quad 0 \leq t \leq T, \ a.s. \tag{3.2}$$

Moreover, in the case : $h(u,e) = -u$*, the inequality* (3.2) *also holds for any triple* $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$ *satisfying* (2.12)*.*

Proof. We state the proof for quadruple $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K})$ satisfying (2.10)-(2.11). Same arguments are used in the case : $h(u,e) = -u$ and $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$ satisfying (2.12).

Denote $\bar{Y} = \tilde{Y} - Y^n$, $\bar{Z} = \tilde{Z} - Z^n$, $\bar{f} = f(X, \tilde{Y}, \tilde{Z}) - f(X, Y^n, Z^n)$ and $\bar{c} = c(X_{.^-}, \tilde{Y}_{.^-}, \tilde{Z}, e) - c(X_{.^-}, Y^n_{.^-}, Z^n, e)$. Fix some $\nu \in \mathcal{V}$ (to be chosen later). We then have :

$$\begin{aligned}\bar{Y}_t = & \int_t^T \bar{f}_s ds + \int_t^T \int_E \bar{c}_s \mu(ds, de) - \int_t^T \langle \bar{Z}_s, dW_s \rangle - \int_t^T \int_E \left\{ \tilde{U}_s(e) - U_s^n(e) \right\} \tilde{\mu}^\nu(ds, de) \\ & - \int_t^T \int_E \left\{ \tilde{U}_s(e) - U_s^n(e) \right\} \nu_s(e) \lambda(de) ds - n \int_t^T \int_E h^-(U_s^n(e), e) \lambda(de) ds + \tilde{K}_T - \tilde{K}_t,\end{aligned}$$

where $\tilde{\mu}^\nu(dt, de) = \mu(dt, de) - \nu_t(e)\lambda(de)dt$ denotes the compensated measure of μ under $\mathbf{P}^\nu$. Let us then define the following adapted processes:

$$a_t \quad = \quad \frac{f(X_t, \tilde{Y}_t, \tilde{Z}_t) - f(X_t, Y_t^n, \tilde{Z}_t)}{\bar{Y}_t} \mathbf{1}_{\{\bar{Y}_t \neq 0\}},$$

and b the $\mathbb{R}^d$-valued process defined by its i-th components, $i = 1, \ldots, d$:

$$b_t^i \quad = \quad \frac{f(X_t, Y_t^n, Z_t^{(i-1)}) - f(X_t, Y_t^n, Z_t^{(i)})}{V_t^i} \mathbf{1}_{\{V_t^i \neq 0\}},$$

where $Z_t^{(i)}$ is the $\mathbb{R}^d$-valued random vector whose i first components are those of $\tilde{Z}$ and whose $(d-i)$ lasts are those of Z^n, and V_t^i is the i-th component of $Z_t^{(i-1)} - Z_t^{(i)}$. Let us also define the $\mathcal{P} \otimes \mathcal{E}$-measurable processes δ in $\mathbb{R}$ and ℓ in $\mathbb{R}^d$ by:

$$\delta_t(e) \quad = \quad \frac{c(X_{t^-}, \tilde{Y}_{t^-}, \tilde{Z}_t) - c(X_{t^-}, Y_{t^-}^n, \tilde{Z}_t, e)}{\bar{Y}_r} \mathbf{1}_{\{\bar{Y}_{t^-} \neq 0\}},$$

and

$$\ell_r^i(e) \quad = \quad \frac{c(X_{t^-}, Y_{t^-}^n, Z_t^{(i-1)}, e) - c(X_{t^-}, Y_{t^-}^n, Z_t^{(i)}, e)}{V_t^i} \mathbf{1}_{\{V_t^i \neq 0\}}.$$

Notice that the processes a, b, δ and ℓ are bounded by the Lipschitz conditions on f and c. Define also $\alpha_t^\nu = a_t + \int_E \delta_t(e)\nu_t(e)\lambda(de)$, $\beta_t^\nu = b_t + \int_E \ell_t(e)\nu_t(e)\lambda(de)$, which are bounded processes since a, b, δ, ℓ are bounded and λ is a finite measure on E, and denote $V_t^n(e) = \tilde{U}_t(e) - U_t^n(e) - \delta_t(e)\bar{Y}_t - \ell_t(e) \cdot \bar{Z}_t$. With these notations, and recalling that $h^-(\tilde{U}_s(e)) = 0$ from the constraint condition (2.11), we rewrite the BSDE for $\bar{Y}$ as:

$$
\begin{aligned}
\bar{Y}_t &= \int_t^T (\alpha_s^\nu \bar{Y}_s + \beta_s^\nu . \bar{Z}_s)ds - \int_t^T \langle \bar{Z}_s, dW_s\rangle - \int_t^T \int_E V_s^n(e)\tilde{\mu}^\nu(ds, de) + \tilde{K}_T - \tilde{K}_t \\
&\quad + \int_t^T \int_E \left\{ n[h^-(\tilde{U}_s(e), e) - h^-(U_s^n(e), e)] - \nu_s(e)[\tilde{U}_s(e) - U_s^n(e)] \right\} \lambda(de)ds.
\end{aligned}
$$

Consider now the positive process Γ^ν solution to the s.d.e.:

$$
d\Gamma_t^\nu \;=\; \Gamma_t^\nu(\alpha_t^\nu dt + \langle \beta_t^\nu, dW_t\rangle), \quad \Gamma_0^\nu \;=\; 1,
$$

and notice that Γ^ν lies in $\mathcal{S}^2$ from the boundeness condition on α^ν and β^ν. By Itô's formula, we have

$$
\begin{aligned}
d\Gamma_t^\nu \bar{Y}_t &= -\Gamma_t^\nu \int_E \left\{ n[h^-(\tilde{U}_t(e), e) - h^-(U_t^n(e), e)] - \nu_t(e)[\tilde{U}_t(e) - U_t^n(e)] \right\} \lambda(de)ds \\
&\quad -\Gamma_t^\nu d\tilde{K}_t + \Gamma_t^\nu \langle \bar{Z}_t, dW_t\rangle + \Gamma_t^\nu \bar{Y}_{t-}\langle \beta_t, dW_t\rangle - \Gamma_t^\nu \int_E V_t^n(e)\tilde{\mu}^\nu(dt, de),
\end{aligned}
$$

which shows that the process

$$
\Gamma_t^\nu \bar{Y}_t + \int_0^t \Gamma_s^\nu \int_E \left\{ n[h^-(\tilde{U}_s(e), e) - h^-(U_s^n(e), e)] - \nu_s(e)[\tilde{U}_s(e) - U_s^n(e)] \right\} \lambda(de)ds
$$

is a $\mathbf{P}^\nu$-supermartingale as soon as $n[h^-(\tilde{U}_t(e), e) - h^-(U_t^n(e), e)] - \nu_t(e)[\tilde{U}_t(e) - U_t^n(e)] \geq 0$ for all $(t, e) \in [0, T] \times E$, and so

$$
\Gamma_t^\nu \bar{Y}_t \geq \mathbf{E}^\nu \left[\int_t^T \Gamma_s^\nu \int_E \left\{ n[h^-(\tilde{U}_s(e), e) - h^-(U_s^n(e), e)] - \nu_s^\varepsilon(e)[\tilde{U}_s(e) - U_s^n(e)] \right\} \lambda(de)ds \,\middle|\, \mathcal{F}_t \right].
$$

Now, from the Lipschitz condition on h, we see that the process ν^ε defined by

$$
\nu_t^\varepsilon(e) \;=\; \begin{cases} \frac{n[h^-(\tilde{U}_s(e), e) - h^-(U_s^n(e), e)]}{\tilde{U}_s(e) - U_s^n(e)} & \text{if } U_t^n(e) > \tilde{U}_s(e) \text{ and } h^-(U_s^n(e), e) > 0 \\ \varepsilon & \text{else} \end{cases}
$$

is bounded and so lies in $\mathcal{V}$, and therefore by taking $\nu = \nu^\varepsilon$, we obtain :

$$
\Gamma_t^{\nu^\varepsilon} \bar{Y}_t \;\geq\; -\varepsilon \mathbf{E}^{\nu^\varepsilon} \left[\int_t^T \Gamma_s^{\nu^\varepsilon} \int_E [\tilde{U}_s(e) - U_s^n(e)] \mathbf{1}_{\{\tilde{U}_s(e) \geq U_s^n(e)\} \cup \{h^-(U_s^n(e), e) = 0\}} \lambda(de)ds \,\middle|\, \mathcal{F}_t \right].
$$

From Bayes formula, this is written as :

$$
\begin{aligned}
&\Gamma_t^{\nu^\varepsilon} \bar{Y}_t \\
\geq\; &-\varepsilon \mathbf{E} \left[\frac{Z_T^{\nu^\varepsilon}}{Z_t^{\nu^\varepsilon}} \int_t^T \Gamma_s^{\nu^\varepsilon} \int_E [\tilde{U}_s(e) - U_s^n(e)] \mathbf{1}_{\{\tilde{U}_s(e) \geq U_s^n(e)\} \cup \{h^-(U_s^n(e), e) = 0\}} \lambda(de)ds \,\middle|\, \mathcal{F}_t \right],
\end{aligned} \tag{3.3}
$$

where Z^{ν^ε} is the Doléans-Dade exponential :

$$Z_t^{\nu^\varepsilon} = \exp\left(-\int_0^t \int_E (\nu_s^\varepsilon(e)-1)\lambda(de)ds\right) \exp\left(\int_0^t \log(\nu_s^\varepsilon(e))\mu(ds,de)\right).$$

By definition of ν^ε, we have

$$\frac{Z_T^{\nu^\varepsilon}}{Z_t^{\nu^\varepsilon}} \leq \frac{Z_T^n}{Z_t^n} \exp\left(\int_t^T (nk_h-1)\lambda(de)ds\right) \exp\left(-\int_t^T \int_E (\varepsilon-1)\lambda(de)ds\right),$$

where Z^n is the solution to $dZ_t^n = Z_{t-}^n \int_E (nk_h-1)\,\tilde{\mu}(dt,de)$, $Z_0^n = 1$. This shows that

$$\mathbf{E}\left[\left(\frac{Z_T^{\nu^\varepsilon}}{Z_t^{\nu^\varepsilon}}\right)^2 \middle| \mathcal{F}_t\right]$$

is uniformly bounded for ε in the neighborhood of 0^+. Notice also that the family $(\Gamma^{\nu^\varepsilon})_{0\leq\varepsilon\leq n}$ is uniformly bounded in $\mathcal{S}^2$ so that

$$\mathbf{E}\left[\left(\int_t^T \Gamma_s^{\nu^\varepsilon} \int_E [\tilde{U}_s(e)-U_s^n(e)]\mathbf{1}_{\{\tilde{U}_s(e)\geq U_s^n(e)\}\cup\{h^-(U_s^n(e),e)=0\}}\lambda(de)ds\right)^2 \middle| \mathcal{F}_t\right]$$

is again uniformly bounded for ε in the neighborhood of 0^+. From the conditional Cauchy-Schwartz inequality, we deduce that

$$\mathbf{E}\left[\frac{Z_T^{\nu^\varepsilon}}{Z_t^{\nu^\varepsilon}} \int_t^T \Gamma_s^{\nu^\varepsilon} \int_E [\tilde{U}_s(e)-U_s^n(e)]\mathbf{1}_{\{\tilde{U}_s(e)\geq U_s^n(e)\}\cup\{h^-(U_s^n(e),e)=0\}}\lambda(de)ds \middle| \mathcal{F}_t\right]$$

is uniformly bounded for ε in the neighborhood of 0^+. Finally, since $\lim_{\varepsilon\to 0}\Gamma_t^{\nu^\varepsilon} = \Gamma_t^{\nu^0} > 0$, by sending ε to zero into (3.3), we conclude that $\bar{Y}_t \geq 0$. □

3.2 Convergence of the penalized BSDEs

We impose the following analogue of Assumption **(H1)**.

(H2) There exists a quadruple $(\tilde{Y},\tilde{Z},\tilde{K},\tilde{U}) \in \mathcal{S}^2 \times \mathbf{L}^2(\mathbf{W}) \times \mathbf{L}^2(\tilde{\mu}) \times \mathbf{A}^2$ satisfying (2.10)-(2.11).

Assumption **(H2)** ensures that the problem (2.10)-(2.11) is well-posed. As indicated in paragraph 2.2, Assumption **(H2)** in the case $h(u,e) = -u$, is stronger than Assumption **(H1)**. We provide in Section 6 some discussion and sufficient conditions under which **(H2)** holds.

Remark 3.1 The following example shows that conditions **(H1)** and **(H2)** may be not satisfied : consider the BSDEs

$$Y_t = -\int_t^T \langle Z_s, dW_s\rangle + \int_t^T \int_E c\mu(ds,de) + K_T - K_t, \tag{3.4}$$

and

$$\begin{cases} Y_t = -\int_t^T \langle Z_s, dW_s\rangle - \int_t^T \int_E [U_s(e)-c]\mu(ds,de) + K_T - K_t \\ -U_s(e) \geq 0 \end{cases} \tag{3.5}$$

where c is a strictly positive constant, $c > 0$. Then, there does not exist any solution to (3.4) or (3.5) with component $Y \in \mathcal{S}^{\mathbf{2}}$. On the contrary, we would have

$$Y_0 \quad \geq \quad -\int_0^T \langle Z_s, dW_s\rangle + c\mu([0,T]\times E), \qquad \text{a.s.}$$

which implies that for all $n \in \mathbb{N}^*$, $\nu \equiv n \in \mathcal{V}$,

$$Y_0 \quad \geq \quad \mathbf{E}^{\nu}\Big[-\int_0^T \langle Z_s, dW_s\rangle + c\mu([0,T]\times E)\Big] \;=\; cn\lambda(E)T.$$

By sending n to infinity, we get the contradiction : $\|Y\|_{\mathcal{S}^2} = \infty$.

We now establish a priori estimates, uniform on n, on the sequence $(Y^n, Z^n, U^n, K^n)_n$.

Lemma 3.3 *Under* **(H2)** *(or* **(H1)** *in the case : $h(u,e) = -u$), there exists some constant C such that*

$$\|Y^n\|_{\mathcal{S}^{\mathbf{2}}} + \|Z^n\|_{\mathbf{L^2(W)}} + \|U^n\|_{\mathbf{L^2}(\tilde\mu)} + \|K^n\|_{\mathcal{S}^{\mathbf{2}}} \quad \leq \quad C, \qquad \forall n \in \mathbb{N}. \tag{3.6}$$

Proof. In what follows we shall denote $C > 0$ to be a generic constant depending only on T, the coefficients f, c, the process X, and the bound for $\tilde Y$ in **(H1)** or **(H2)**, and which may vary from line to line.

Applying Itô's formula to $|Y_t^n|^2$, and observing that K^n is continuous and $\Delta Y_t^n = \int_E U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)\mu(\{t\}, de)$, we have

$$\begin{aligned}\mathbf{E}|g(X_T)|^2 = \mathbf{E}|Y_t^n|^2 - 2\mathbf{E}\int_t^T Y_s^n f(X_s, Y_s^n, Z_s^n)ds - 2\mathbf{E}\int_t^T Y_s^n dK_s^n + \mathbf{E}\int_t^T |Z_s^n|^2 ds \\ + \mathbf{E}\int_t^T\int_E \big\{|Y_{s-}^n + U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)|^2 - |Y_{s-}^n|^2\big\}\lambda(de)ds\end{aligned}$$

From the linear growth condition on f and the inequality $Y_t^n \leq \tilde Y_t$ by Lemma 3.2 under **(H2)** (and also under **(H1)** in the case $h(u,e) = -u$), and using the inequality $2ab \leq \frac{1}{\alpha}a^2 + \alpha b^2$ for any constant $\alpha > 0$, we have:

$$\begin{aligned}&\mathbf{E}|Y_t^n|^2 + \mathbf{E}\int_t^T |Z_s^n|^2 ds + \mathbf{E}\int_t^T\int_E |U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)|^2\lambda(de)ds \\ \leq\;& \mathbf{E}|g(X_T)|^2 + 2C\mathbf{E}\int_t^T |Y_s^n|\,(1 + |X_s| + |Y_s^n| + |Z_s^n|)\,ds \\ &-2\mathbf{E}\int_t^T\int_E Y_{s-}^n(U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e))\lambda(de)ds + \frac{1}{\alpha}\mathbf{E}\Big[\sup_{t\in[0,T]}|\tilde Y_t|^2\Big] + \alpha\mathbf{E}|K_T^n - K_t^n|^2.\end{aligned}$$

Using again the inequality $2ab \leq \frac{1}{\alpha}a^2 + \alpha b^2$, in particular for $\alpha = 2$, yields

$$\begin{aligned}&\mathbf{E}|Y_t^n|^2 + \mathbf{E}\int_t^T |Z_s^n|^2 ds + \frac{1}{2}\mathbf{E}\int_t^T\int_E |U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)|^2\lambda(de)ds \\ \leq\;& \mathbf{E}|g(X_T)|^2 + 2C\mathbf{E}\int_t^T |Y_s^n|\,(1 + |X_s| + |Y_s^n| + |Z_s^n|)\,ds \\ &+ 2\lambda(E)\mathbf{E}\int_t^T |Y_s^n|^2 ds + \frac{1}{\alpha}\mathbf{E}\Big[\sup_{t\in[0,T]}|\tilde Y_t|^2\Big] + \alpha\mathbf{E}|K_T^n - K_t^n|^2 \\ \leq\;& C\Big(1 + \mathbf{E}\int_t^T |Y_s^n|^2 ds\Big) + \frac{1}{2}\mathbf{E}\int_t^T |Z_s^n|^2 ds + \alpha\mathbf{E}|K_T^n - K_t^n|^2.\end{aligned}$$

Then, by using the inequality $(a-b)^2 \geq a^2/2 - b^2$, we get

$$\begin{aligned}
& \mathbf{E}|Y_t^n|^2 + \frac{1}{2}\mathbf{E}\int_t^T |Z_s^n|^2 ds + \frac{1}{4}\mathbf{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de)ds \\
\leq\ & \frac{1}{2}\mathbf{E}\int_t^T \int_E |c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)|^2 \lambda(de)ds \\
& + C\Big(1 + \mathbf{E}\int_t^T |Y_s^n|^2 ds\Big) + \alpha \mathbf{E}|K_T^n - K_t^n|^2 \\
\leq\ & C\Big(1 + \mathbf{E}\int_t^T |Y_s^n|^2 ds\Big) + \frac{1}{4}\mathbf{E}\int_t^T |Z_s^n|^2 ds + \alpha \mathbf{E}|K_T^n - K_t^n|^2, \qquad (3.7)
\end{aligned}$$

from the linear growth condition on c. Now, from the relation

$$\begin{aligned}
K_T^n - K_t^n \ =\ & Y_t^n - g(X_T) - \int_t^T f(X_s, Y_s^n, Z_s^n) ds \\
& + \int_t^T \int_E (U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n))\mu(ds, de) + \int_t^T Z_s^n . dW_s,
\end{aligned}$$

and the linear growth condition on f, c, there exists some positive constant C_1 s.t.

$$\begin{aligned}
& \mathbf{E}|K_T^n - K_t^n|^2 \\
\leq\ & C_1 \left(1 + \mathbf{E}|Y_t^n|^2 + \mathbf{E}\int_t^T |Y_s^n|^2 + |Z_s^n|^2 ds + \mathbf{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de) ds\right). \qquad (3.8)
\end{aligned}$$

Hence, by choosing $\alpha > 0$ s.t. $C_1\alpha < 1/4$, and plugging into (3.7), we get

$$\mathbf{E}|Y_t^n|^2 + \mathbf{E}\int_t^T |Z_s^n|^2 ds + \mathbf{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de) ds \ \leq\ C\left(1 + \mathbf{E}\int_t^T |Y_s^n|^2 ds\right).$$

By applying Gronwall's lemma to $t \mapsto \mathbf{E}|Y_t^n|^2$ and (3.8), we obtain

$$\sup_{0\leq t\leq T} \mathbf{E}|Y_t^n|^2 + \mathbf{E}\int_0^T |Z_s^n|^2 ds + \mathbf{E}\int_0^T \int_E |U_s^n(e)|^2 \lambda(de) ds + \mathbf{E}|K_T^n|^2 \ \leq\ C. \qquad (3.9)$$

Finally, by writing from (3.1) that

$$\begin{aligned}
\sup_{0\leq t\leq T} |Y_t^n| \ \leq\ & |g(X_T)| + \int_0^T |f(X_s, Y_s, Z_s)| ds + K_T^n \\
& + \sup_{s\in[0,T]} \left|\int_0^T \langle Z_s, dW_s\rangle\right| + \int_0^T \int_E |U_s^n(e) - c(X_{s^-}, Y_{s^-}, Z_s, e)|\mu(ds, de),
\end{aligned}$$

we obtain the required result from the Burkholder-Davis-Gundy inequality, the linear growth condition on f, c, and (3.9). □

Remark 3.2 A closer look at the proof leading to the estimate in (3.6) shows that there exists a universal constant C, depending only on T, and the linear growth condition constants of f, c, such that for each $n \in \mathbb{N}$:

$$\sup_{t\in[0,T]} \mathbf{E}[Y_t^n]^2 \ \leq\ C\Big(1 + \mathbf{E}|g(X_T)|^2 + \mathbf{E}\Big[\int_0^T |X_t|^2 dt\Big] + \mathbf{E}\Big[\sup_{t\in[0,T]} |\tilde{Y}_t|^2\Big]\Big). \qquad (3.10)$$

Lemma 3.4 *Under* **(H2)** *(or* **(H1)** *in the case : $h(u,e) = -u$), the sequence of processes (Y_t^n) converges increasingly to a process (Y_t) with $Y \in \mathcal{S}^2$. The convergence also holds in $\mathbf{L}^2_{\mathbb{F}}(\mathbf{0},\mathbf{T})$ and for every stopping time $\tau \in [0,T]$, the sequence of random variables (Y_τ^n) converges to Y_τ in $\mathbf{L}^2(\boldsymbol{\Omega},\mathcal{F}_\tau)$, i.e.*

$$\lim_{n\to\infty} \mathbf{E}\Big[\int_0^T |Y_t^n - Y_t|^2 dt\Big] = 0 \quad \textit{and} \quad \lim_{n\to\infty} \mathbf{E}\Big[|Y_\tau^n - Y_\tau|^2\Big] = 0. \tag{3.11}$$

Proof. From Lemmas 3.1 and 3.2, the (nondecreasing) limit

$$Y_t \;\; := \;\; \lim_{n\to\infty} Y_t^n, \quad 0 \le t \le T, \tag{3.12}$$

exists almost surely, and this defines an adapted process Y. Moreover, by Lemma 3.3 and convergence monotone theorem, we have

$$\mathbf{E}\Big[\sup_{0\le t\le T} |Y_t|^2\Big] \;\; < \;\; \infty.$$

From the dominated convergence theorem, we also get the convergences (3.11). It remains to check that the process Y has a càdlàg modification. We first show that $(Y^n)_n$ are quasi-martingales with uniformly bounded conditional variations. That is, there exists a constant C such that, for any partition $\pi : 0 = t_0 < t_1 < \cdots < t_n = T$,

$$\mathbf{E}\Big\{|Y_T^n| + \sum_{i=0}^{n-1} |\mathbf{E}\{Y_{t_{i+1}}^n|\mathcal{F}_{t_i}\} - Y_{t_i}^n|\Big\} \le C, \quad \forall \pi,\ \forall n. \tag{3.13}$$

In fact, by (3.1) we have

$$\begin{aligned}
&\mathbf{E}\Big\{\sum_{i=0}^{n-1} |\mathbf{E}\{Y_{t_{i+1}}^n|\mathcal{F}_{t_i}\} - Y_{t_i}^n|\Big\} = \mathbf{E}\Big\{\sum_{i=0}^{n-1}\Big|\mathbf{E}\Big[\int_{t_i}^{t_{i+1}} f(X_s, Y_s^n, Z_s^n)ds \\
&+n\int_{t_i}^{t_{i+1}}\int_E h^-(U_s^n(e),e)\lambda(de)ds - \int_{t_i}^{t_{i+1}}\int_E (U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e))\lambda(de)ds\Big|\mathcal{F}_{t_i}\Big]\Big|\Big\} \\
&\le \mathbf{E}\Big[\int_0^T |f(X_s, Y_s^n, Z_s^n)|ds + \int_0^T\int_E |U_s^n(e) - c(X_{s^-}, Y_{s^-}^n, Z_s^n, e)|\lambda(de)ds + K_T^n\Big].
\end{aligned}$$

Recall (2.3), (2.4), and (2.6), we have

$$\begin{aligned}
&\mathbf{E}\Big\{|Y_T^n| + \sum_{i=0}^{n-1} |\mathbf{E}\{Y_{t_{i+1}}^n|\mathcal{F}_{t_i}\} - Y_{t_i}^n|\Big\} \\
&\le C\mathbf{E}\Big\{1 + |X_T| + \int_0^T [1 + |X_s| + |Y_s^n| + |Z_s^n|]ds + \int_0^T\int_E |U_s^n(e)|\lambda(de)ds + K_T^n\Big\}.
\end{aligned}$$

Applying (2.2) and Lemma 3.3, we obtain (3.13) immediately. Now by Meyer-Zheng [15] (or see [14]), there exists a subsequence $(Y^{n_k})_k$ and a càdlàg process $\tilde{Y}$ such that $(Y^{n_k})_k$ converges to $\tilde{Y}$ in distribution. On the other hand, by (3.12), $(Y^{n_k})_k$ converges to Y, P-a.s. Then Y and $\tilde{Y}$ have the same distribution, and thus Y is also càdlàg. □

We now focus on the convergence of the diffusion and jump components (Z^n, U^n). In our context, we cannot prove the strong convergence of (Z^n, U^n) in $\mathbf{L}^2(\mathbf{W}) \times \mathbf{L}^2(\tilde{\mu})$, and so the strong convergence of $\int_0^t Z^n dW$ and $\int_0^t \int_E U^n(s,e)\mu(ds,de)$ in $\mathbf{L^2(\Omega,\mathcal{F}_t)}$, see Remark 3.3. Instead, we follow and extend arguments of Peng [18], and we shall prove that (Z^n, U^n) converge in $\mathbf{L^p(W)} \times \mathbf{L^p}(\tilde{\mu})$, for $1 \le p < 2$. First, we show the following weak convergence and decomposition result.

Lemma 3.5 *Under* **(H2)** *(or* **(H1)** *in the case: $h(u,e) = -u$), there exist $\phi \in \mathbf{L^2_{\mathbb{F}}(0,T)}$, $Z \in \mathbf{L^2(W)}$, $V \in \mathbf{L^2}(\tilde{\mu})$ and $K \in \mathbf{A^2}$ predictable, such that the limit Y in* (3.12) *has the form*

$$Y_t \;=\; Y_0 - \int_0^t \phi_s ds - K_t + \int_0^t \langle Z_s, dW_s\rangle + \int_0^t \int_E V_s(e)\mu(ds,de), \quad 0 \le t \le T. \tag{3.14}$$

Moreover, in the above decomposition of Y, the components Z and V are unique, and are respectively the weak limits of (Z^n) in $\mathbf{L^2}(\tilde{\mu})$ and of (V^n) in $\mathbf{L^2}(\tilde{\mu})$ where $V_t^n(e) = U_t^n(e) - c(X_{t^-}, Y_{t^-}^n, Z_t^n, e)$, ϕ is the weak limit in $\mathbf{L^2_{\mathbb{F}}(0,T)}$ of a subsequence of $(f^n) := (f(X,Y^n,Z^n))$, and K is the weak limit in $\mathbf{L^2_{\mathbb{F}}(0,T)}$ of a subsequence of (K^n).

Proof. By Lemma 3.3, and the linear growth conditions on f, c together with (2.2), the sequences (f^n), (Z^n), (V^n) are weakly compact, respectively in $\mathbf{L^2_{\mathbb{F}}(0,T)}$, $\mathbf{L^2(W)}$ and $\mathbf{L^2}(\tilde{\mu})$. Then, up to a subsequence, (f^n), (Z^n), (V^n) converge weakly to ϕ, Z and V. By Itô representation of martingales, we then get the following weak convergence in $\mathbf{L^2}(\mathbf{\Omega},\mathcal{F}_\tau)$ for each stopping time $\tau \le T$:

$$\int_0^\tau f_s^n ds \rightharpoonup \int_0^\tau \phi_s ds, \qquad \int_0^\tau \langle Z_s^n, dW_s\rangle \rightharpoonup \int_0^\tau \langle Z_s, dW_s\rangle,$$
$$\int_0^\tau \int_E V_s^n(e)\mu(ds,de) \rightharpoonup \int_0^\tau \int_E V_s(e)\mu(ds,de).$$

Since, we have from (3.1):

$$K_\tau^n \;=\; -Y_\tau^n + Y_0^n - \int_0^\tau f_s^n ds + \int_0^\tau \langle Z_s^n, dW_s\rangle + \int_0^\tau \int_E V_s^n(e)\mu(ds,de), \tag{3.15}$$

we also have the weak convergence in $\mathbf{L^2}(\mathbf{\Omega},\mathcal{F}_\tau)$:

$$K_\tau^n \;\rightharpoonup\; K_\tau \;:=\; -Y_\tau + Y_0 - \int_0^\tau \phi_s ds + \int_0^\tau \langle Z_s, dW_s\rangle + \int_0^\tau \int_E V_s(e)\mu(ds,de). \tag{3.16}$$

The process K inherits from K^n the nondecreasing path property, is square integrable, càd-làg and adapted from (3.16), and so lies in $\mathbf{A^2}$. Moreover, by dominated convergence theorem, we see that K^n converges weakly to K in $\mathbf{L^2(0,T)}$. Since K^n is continuous, and so predictable, we deduce that K is also predictable, and we obtain the decomposition (3.14) for Y. The uniqueness of Z follows by identifying the Brownian parts and finite variation parts, and the uniqueness of V is then obtained by identifying the predictable parts and by recalling that the jumps of μ are inacessible. We conclude that (Z,V) is uniquely determined in (3.14), and thus the whole sequence (Z^n, V^n) converges weakly to (Z,V) in $\mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu})$. □

The sequence (U^n) is bounded in $\mathbf{L^2}(\tilde{\mu})$, and so, up to a subsequence, converges weakly to some $U \in \mathbf{L^2}(\tilde{\mu})$. The next step is to show that the whole sequence (U^n) converges to U and to identify in the decomposition (3.14) ϕ_t with $f(X_t, Y_t, Z_t)$, and $V_t(e)$ with $U_t(e) - c(X_{t^-}, Y_{t^-}, Z_t, e)$. Since f and c are nonlinear, we need a result of strong convergence for (Z^n) and (U^n) to enable us to pass the limit in $f(X_t, Y_t^n, Z_t^n)$ as well as in $U_t^n(e) - c(X_{t^-}, Y_{t^-}^n, Z_t^n, e)$, and to eventually prove the convergence of the penalized BSDEs to the minimal solution of our jump-constrained BSDE. We shall borrow a useful technique of Peng [18] to carry out this task.

Theorem 3.1 *Under* **(H2)**, *there exists a unique minimal solution* $(Y, Z, U, K) \in \mathcal{S}^2 \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ *with K predictable, to* (2.10)-(2.11)*. Y is the increasing limit of (Y^n) in* (3.12) *and also in $\mathbf{L^2_{\mathbb{F}}(0,T)}$, K is the weak limit of (K^n) in $\mathbf{L^2_{\mathbb{F}}(0,T)}$, and for any $p \in [1,2)$,*

$$\|Z^n - Z\|_{\mathbf{L^p(W)}} + \|U^n - U\|_{\mathbf{L^p}(\tilde{\mu})} \quad \longrightarrow \quad 0,$$

as n goes to infinity. Moreover, in the case : $h(u,e) = -u$, $(Y, Z, \bar{K})$ is the unique minimal solution to (2.12) *with $\bar{K}_t = K_t - \int_0^t \int_E U_s(e)\mu(ds, de)$, and this holds true under* **(H1)**. *Consequently, the minimal solution Y to* (2.12) *and to* (2.10)-(2.13) *are the same.*

Proof. We apply Itô's formula to $|Y_t^n - Y_t|^2$ on a subinterval $(\sigma, \tau]$, with $0 \le \sigma < \tau \le T$, two stopping times. Recall the decomposition (3.14), (3.15) of Y, Y^n, and observe that K^n is continuous, and $\Delta(Y_t^n - Y_t) = \Delta K_t + \int_E (V_s^n(e) - V_s(e))\mu(\{t\}, de)$. We then have :

$$\begin{aligned}
\mathbf{E}|Y_\tau^n - Y_\tau|^2 \;=\;& \mathbf{E}|Y_\sigma^n - Y_\sigma|^2 + \mathbf{E}\int_\sigma^\tau |Z_s^n - Z_s|^2 ds + 2\mathbf{E}\int_\sigma^\tau [Y_s^n - Y_s][\phi_s - f_s^n] ds \\
& - 2\mathbf{E}\int_\sigma^\tau [Y_s^n - Y_s] dK_s^n + 2\mathbf{E}\int_{(\sigma,\tau]} [Y_{s^-}^n - Y_{s^-}] dK_s + \mathbf{E}\sum_{t\in(\sigma,\tau]} |\Delta K_t|^2 \\
& + \mathbf{E}\int_{(\sigma,\tau]}\int_E [|Y_{s^-}^n - Y_{s^-} + V_s^n(e) - V_s(e)|^2 - |Y_{s^-}^n - Y_{s^-}|^2]\mu(ds, de) \\
=\;& \mathbf{E}|Y_\sigma^n - Y_\sigma|^2 + \mathbf{E}\int_\sigma^\tau |Z_s^n - Z_s|^2 ds + 2\mathbf{E}\int_\sigma^\tau [Y_s^n - Y_s][\phi_s - f_s^n] ds \\
& - 2\mathbf{E}\int_\sigma^\tau [Y_s^n - Y_s] dK_s^n \; + \; 2\mathbf{E}\int_{(\sigma,\tau]} [Y_{s^-}^n - Y_{s^-} + \Delta K_s] dK_s \\
& - \mathbf{E}\sum_{t\in(\sigma,\tau]} |\Delta K_t|^2 + \mathbf{E}\int_\sigma^\tau \int_E |V_s^n(e) - V_s(e)|^2 \lambda(de) ds \\
& + 2\mathbf{E}\int_\sigma^\tau \int_E (Y_s^n - Y_s)(V_s^n(e) - V_s(e))\lambda(de) ds.
\end{aligned}$$

Since $(Y_s^n - Y_s) dK_s^n \le 0$, and by using the inequality $2ab \ge -\frac{a^2}{2} - 2b^2$ with $a = V_s^n(e) - V_s(e)$ and $b = Y_s^n - Y_s$, we obtain :

$$\begin{aligned}
& \mathbf{E}\int_\sigma^\tau |Z_s^n - Z_s|^2 ds + \frac{1}{2}\mathbf{E}\int_\sigma^\tau \int_E |V_s^n(e) - V_s(e)|^2 ds \\
\le\;& \mathbf{E}|Y_\tau^n - Y_\tau|^2 + 2\mathbf{E}\int_\sigma^\tau |Y_s^n - Y_s|^2 ds + 2\mathbf{E}\int_\sigma^\tau |Y_s^n - Y_s||\phi_s - f_s^n| ds \\
& + 2\mathbf{E}\int_{(\sigma,\tau]} |Y_{s^-}^n - Y_{s^-} + \Delta K_s| dK_s + \mathbf{E}\sum_{t\in(\sigma,\tau]} |\Delta K_t|^2. \qquad (3.17)
\end{aligned}$$

The two first terms of the right side of (3.17) converge to zero by (3.11) in Lemma 3.4. The third term also tends to zero since $(\phi - f^n)_n$ is bounded in $\mathbf{L^2(0,T)}$, and so by Cauchy-Schwarz inequality:

$$\mathbf{E}\int_0^T |Y_s^n - Y_s||\phi_s - f_s^n|ds \;\; \leq \;\; C\Big(\mathbf{E}\int_0^T |Y_s^n - Y_s|^2 ds\Big)^{\frac{1}{2}} \;\to\; 0. \qquad (3.18)$$

For the fourth term, we notice that the jumps of Y^n are inacessible since they are determined by the Poisson random measure μ. Thus, the predictable projection of Y^n is ${}^pY_t^n = Y_{t^-}^n$. Similarly, from (3.14), and since K is predictable, we see that ${}^pY_t = Y_{t^-} - \Delta K_t$. Since Y^n increasingly converges to Y, then ${}^pY^n$ also increasingly converges to ${}^pY$, and by the dominated convergence theorem, we obtain:

$$\lim_{n\to\infty} \mathbf{E}\int_{(0,T]} |Y_{s^-}^n - Y_{s^-} + \Delta K_s| dK_s \;\; = \;\; 0. \qquad (3.19)$$

For the last term in (3.17), we apply Lemma 2.3 in [18] to the predictable nondecreasing process K: for any $\delta, \varepsilon > 0$, there exist a finite number of pairs of stopping times (σ_k, τ_k), $k = 0, \ldots, N$, with $0 < \sigma_k \leq \tau_k \leq T$, such that all the intervals $(\sigma_k, \tau_k]$ are disjoint and

$$\mathbf{E}\sum_{k=0}^N (\tau_k - \sigma_k) \geq T - \frac{\varepsilon}{2}, \qquad \mathbf{E}\sum_{k=0}^N \sum_{\sigma_k < t \leq \tau_k} (\Delta K_t)^2 \leq \frac{\varepsilon\delta}{3}. \qquad (3.20)$$

We should note that in [18] the filtration is Brownian, therefore it is continuous, and hence each stopping time σ_k can be approximated by a sequence of announceable stopping times In our case the stopping times σ_k's are constructed as the successive times of jumps of the predictable process K with size bigger than some given positive level, the approximation of σ_k by announceable stopping times is again possible. We can thus argue exactly the same way as in Lemma 2.3 in [18] to derive both estimates in (3.20).

We now apply estimate (3.17) for each $\sigma = \sigma_k$ and $\tau = \tau_k$, and then take the sum over $k = 0, \ldots, N$. It follows that

$$\begin{aligned} & \sum_{k=0}^N \mathbf{E}\int_{\sigma_k}^{\tau_k} |Z_s^n - Z_s|^2 ds + \frac{1}{2}\sum_{k=0}^N \mathbf{E}\int_{\sigma_k}^{\tau_k}\int_E |V_s^n(e) - V_s(e)|^2 ds \\ \leq \;\; & \sum_{k=0}^N \mathbf{E}|Y_{\tau_k}^n - Y_{\tau_k}|^2 + 2\mathbf{E}\int_0^T |Y_s^n - Y_s|^2 ds + 2\mathbf{E}\int_0^T |Y_s^n - Y_s||\phi_s - f_s^n| ds \\ & + 2\mathbf{E}\int_{(0,T]} |Y_{s^-}^n - Y_{s^-} + \Delta K_s| dK_s + \sum_{k=0}^N \mathbf{E}\sum_{t\in(\sigma_k,\tau_k]} |\Delta K_t|^2. \end{aligned}$$

From the convergence results in Lemma 3.4, (3.18) and (3.19), we deduce that

$$\begin{aligned} & \limsup_{n\to\infty} \sum_{k=0}^N \mathbf{E}\int_{\sigma_k}^{\tau_k} |Z_s^n - Z_s|^2 ds + \frac{1}{2}\sum_{k=0}^N \mathbf{E}\int_{\sigma_k}^{\tau_k}\int_E |V_s^n(e) - V_s(e)|^2 \lambda(de) ds \\ \leq \;\; & \sum_{k=0}^N \mathbf{E}\sum_{t\in(\sigma_k,\tau_k]} |\Delta K_t|^2 \;\; \leq \;\; \frac{\varepsilon\delta}{3}. \end{aligned}$$

Thus, there exists an integer $\ell_{\varepsilon\delta} > 0$ such that for all $n \geq \ell_{\varepsilon\delta}$, we have

$$\sum_{k=0}^{N} \mathbf{E} \int_{\sigma_k}^{\tau_k} |Z_s^n - Z_s|^2 ds + \frac{1}{2} \sum_{k=0}^{N} \mathbf{E} \int_{\sigma_k}^{\tau_k} \int_E |V_s^n(e) - V_s(e)|^2 \lambda(de) ds \quad \leq \quad \frac{\varepsilon\delta}{2}.$$

This implies

$$dt \otimes \mathbf{P}\Big[(s,\omega) \in \bigcup_{k=0}^{N} (\sigma_k(\omega), \tau_k(\omega)] \times \Omega \ : |Z_s^n(\omega) - Z_s(\omega)|^2 \geq \delta\Big] \quad \leq \quad \frac{\varepsilon}{2},$$

and

$$dt \otimes \lambda \otimes \mathbf{P}\Big[(s,e,\omega) \in \bigcup_{k=0}^{N} (\sigma_k(\omega), \tau_k(\omega)] \times \Omega \times E \ : |V_s^n(e,\omega) - V_s(e,\omega)|^2 \geq \delta\Big] \quad \leq \quad \varepsilon.$$

Together with (3.20), it follows that $dt \otimes \mathbf{P}\left[(s,\omega) \in [0,T] \times \Omega \ : |Z_s^n(\omega) - Z_s(\omega)|^2 \geq \delta\right] \leq \varepsilon$, and

$$dt \otimes \lambda \times \mathbf{P}\left[(s,e,\omega) \in [0,T] \times E \times \Omega \ : |V_s^n(e,\omega) - V_s(e,\omega)|^2 \geq \delta\right] \quad \leq \quad \varepsilon(1 + \lambda(E)).$$

We deduce that for all $\delta > 0$,

$$\lim_{n \to \infty} dt \otimes \mathbf{P}\left[(s,\omega) \in [0,T] \times \Omega \ : |Z_s^n(\omega) - Z_s(\omega)|^2 \geq \delta\right] \quad = \quad 0$$

and

$$\lim_{n \to \infty} dt \otimes \lambda \otimes \mathbf{P}\left[(s,e,\omega) \in [0,T] \times E \times \Omega \ : |V_s^n(e,\omega) - V_s(e,\omega)|^2 \geq \delta\right] \quad = \quad 0.$$

This means that the sequences $(Z^n)_n$ and $(V^n)_n$ converge in measure respectively to Z and V. Since they are bounded respectively in $\mathbf{L^2(W)}$ and $\mathbf{L^2}(\tilde{\mu})$, they are uniformly integrable in $\mathbf{L^p(W)}$ and $\mathbf{L^p}(\tilde{\mu})$ for any $p \in [1,2)$, respectively. Thus, (Z^n) and (V^n) converge strongly to Z and V in $\mathbf{L^p(W)}$ and $\mathbf{L^p}(\tilde{\mu})$, respectively. Recalling that $U_t^n(e) = V_t^n(e) + c(X_{t-}, Y_{t-}^n, Z_t^n, e)$, and by the Lipschitz condition on c, we deduce that the sequence (U^n) converges strongly in $\mathbf{L^p}(\tilde{\mu})$, for $p \in [1,2)$, to U defined by :

$$U_t(e) \quad = \quad V_t(e) + c(X_{t-}, Y_{t-}, Z_t, e), \quad 0 \leq t \leq T, \ e \in E.$$

By the Lipschitz condition on f, we also have the strong convergence in $\mathbf{L^p_{\mathbb{F}}(0,T)}$ of $(f^n) = (f(X, Y^n, Z^n))$ to $f(X,Y,Z)$. Since ϕ is the weak limit of (f^n) in $\mathbf{L^2_{\mathbb{F}}(0,T)}$, we deduce that $\phi = f(X,Y,Z)$. Therefore, with the decomposition (3.14) and since $Y_T = \lim_n Y_T^n = g(X_T)$, we obtain immediately that (Y,Z,U,K) satisfies the BSDE (2.10). Moreover, from the strong convergence in $\mathbf{L^1}(\tilde{\mu})$ of (U^n) to U, and the Lipschitz condition on h, we have

$$\mathbf{E} \int_0^T \int_E h^-(U_s^n(e), e) \lambda(de) ds \quad \to \quad \mathbf{E} \int_0^T \int_E h^-(U_s(e), e) \lambda(de) ds,$$

as n goes to infinity. Since $K_T^n = n \int_0^T \int_E h^-(U_s^n(e), e) \lambda(de) ds$ is bounded in $\mathbf{L^2(\Omega, \mathcal{F}_T)}$, this implies

$$\mathbf{E} \int_0^T \int_E h^-(U_s(e), e) \lambda(de) ds \quad = \quad 0,$$

and so the constraint (2.11) is satisfied. Hence, (Y, Z, K, U) is a solution to the constrained BSDE (2.10)-(2.11), and by Lemma 3.2, $Y = \lim Y^n$ is the minimal solution. The uniqueness of Z follows by identifying the Brownian parts and the finite variation parts, and then the uniqueness of (U, K) is obtained by identifying the predictable parts and by recalling that the jumps of μ are inacessible.

Finally, in the case $h(u, e) = -u$, the process

$$\bar{K}_t \;=\; K_t - \int_0^t \int_E U_s(e)\mu(ds, de), \quad 0 \le t \le T,$$

lies in $\mathbf{A^2}$, and the triple $(Y, Z, \bar{K})$ is solution to (2.12). Again, by Lemma 3.2, this shows that Y is the minimal solution to (2.10) and to (2.12). The uniqueness of $(Y, Z, \bar{K})$ is immediate by identifying the Brownian part and the finite variation part. □

Remark 3.3 From the estimate (3.17), it is clear that once the process K is continuous, i.e. $\Delta K_t = 0$, then (Z^n, U^n) converges strongly to (Z, U) in $\mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu})$. This occurs in reflected BSDE's as in [10] or [12], see also Remark 4.3. In the case of constraints on jump component U as in (2.10)-(2.11), the situation is more complicated, and the process K is in general only predictable. The same feature also occurs for constraints on Z as in [18]. To overcome this difficulty, we use the estimations (3.20) of the contribution of the jumps of K, which allow to obtain the strong convergence of (Z^n, U^n) in $\mathbf{L^p(W)} \times \mathbf{L^p}(\tilde{\mu})$ for $p \in [1, 2)$. Finally, notice that for the minimal solution $(Y, Z, \tilde{K})$ to the BSDE (2.12), the process $\tilde{K}$ is not predictable.

3.3 The case of impulse control

In the impulse control case (i.e. f and c depend only on X and $h(u, e) = -u$), we have seen in Theorem 2.1 that the minimal solution to our constrained BSDE has the following functional explicit representation :

$$Y_t \;=\; \operatorname*{ess\,sup}_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds, de) \Big| \mathcal{F}_t \Big].$$

In this case, we also have a functional explicit representation of the solution Y^n to the penalized BSDE (3.1) :

$$Y_t^n \;=\; \operatorname*{ess\,sup}_{\nu \in \mathcal{V}_n} \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-})\mu(ds, de) \Big| \mathcal{F}_t \Big], \tag{3.21}$$

where $\mathcal{V}_n = \{\nu \in \mathcal{V} \; ; \; \nu_s(e) \le n \; \forall (s, e) \in [0, T] \times E \; a.s.\}$. Indeed, denote by $\bar{Y}^n$ the right side of (1.4). By writing that (Y^n, Z^n, U^n) is the solution of the penalized BSDE (3.1), taking the expectation under $\mathbf{P}^\nu$, for $\nu \in \mathcal{V}_n$, and recalling that W is a $\mathbf{P}^\nu$-Brownian motion, and $\nu\lambda(de)$ is the intensity measure of μ under $\mathbf{P}^\nu$, we obtain :

$$\begin{aligned} Y_t^n \;=\;& \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds, de) \Big| \mathcal{F}_t \Big] \\ &+ \mathbf{E}^\nu \Big[\int_t^T \int_E \{ n[U_s^n(e)]_+ - \nu_s(e) U_s^n(e) \} \, \lambda(de)ds \Big| \mathcal{F}_t \Big]. \end{aligned} \tag{3.22}$$

Since this equality holds for any $\nu \in \mathcal{V}_n$, and observing that $n[U_s^n(e)]_+ - \nu_s(e)U_s^n(e) \geq 0$, for all $\nu \in \mathcal{V}_n$, we have

$$\bar{Y}_t^n \ \leq \ Y_t^n \ \ \leq \ \ \tilde{Y}_t^n + \mathbf{E}^\nu\Big[\int_t^T \int_E \{n[U_s^n(e)]_+ - \nu_s(e)U_s^n(e)\}\,\lambda(de)ds\Big|\mathcal{F}_t\Big]. \qquad (3.23)$$

Let us now consider the family $(\nu^\varepsilon)_\varepsilon$ of $\mathcal{V}_n$ defined by

$$\nu_s^\varepsilon(e) \quad = \quad \begin{cases} \frac{n[U_s^n(e)]_+}{U_s^n(e)} & \text{if } U_s^n(e) > 0 \\ \varepsilon & \text{otherwise.} \end{cases}$$

Then, by using the same argument as in the proof of Lemma 3.2, we show that

$$\mathbf{E}^{\nu^\varepsilon}\Big[\int_t^T \int_E \{n[U_s^n(e)]_+ - \nu_s(e)U_s^n(e)\}\,\lambda(de)ds\Big|\mathcal{F}_t\Big] \quad \rightarrow \quad 0 \text{ as } \varepsilon \rightarrow 0,$$

which proves with (3.23) that $Y_t^n = \bar{Y}_t^n$.

The representation (3.21) has a nice interpretation. It means that the value function of an impulse control problem can be approximated by the value function of the same impulse control problem but with strategies whose numbers of orders are bounded *on average* by $nT\lambda(E)$. This has to be compared with the classical approximation by iterated optimal stopping problems, where the n-th iteration corresponds to the value of the same impulse control problem but where the number of orders is smaller than n. The numerical advantage of the penalized approximation is that it does not require iterations.

4 Relation with quasi-variational inequalities

In this section, we show that minimal solutions to the jump-constrained BSDEs provide a probabilistic representation of solutions to parabolic QVIs of the form:

$$\min\Big[-\frac{\partial v}{\partial t} - \mathcal{L}v - f(\cdot, v, \sigma^\intercal D_x v)\ ,\ \inf_{e\in E} h(\mathcal{H}^e v - v, e)\Big] \quad = \quad 0, \quad \text{ on } [0,T) \times \mathbb{R}^d, \ (4.1)$$

where $\mathcal{L}$ is the second order local operator

$$\mathcal{L}v(t,x) = \langle b(x), D_x v(t,x)\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x)D_x^2 v(t,x)),$$

and $\mathcal{H}^e$, $e \in E$, are the nonlocal operators

$$\mathcal{H}^e v(t,x) = v(t, x + \gamma(x,e)) + c(x, v(t,x), \sigma^\intercal(x)D_x v(t,x), e).$$

For such nonlocal operators, we denote for $q \in \mathbb{R}^d$:

$$\mathcal{H}^e[t,x,q,v] \quad = \quad v(t, x+\gamma(x,e)) + c(x, v(t,x), \sigma^\intercal(x)q, e).$$

Note that when $h(u)$ does not depend on e, and since it is nonincreasing in u, the QVI (4.1) may be written equivalently in

$$\min\Big[-\frac{\partial v}{\partial t} - \mathcal{L}v - f(\cdot, v, \sigma^\intercal D_x v)\ ,\ h(\mathcal{H}v - v)\Big] \quad = \quad 0, \quad \text{ on } [0,T) \times \mathbb{R}^d,$$

with $\mathcal{H}v = \sup_{e\in E}\mathcal{H}^e v$. In particular, this includes the case of QVI associated to impulse controls for $h(u) = -u$, and f, c independent of y, z.

We shall use the penalized parabolic integral partial differential equation (IPDE) associated to the penalized BSDE (3.1), for each $n \in \mathbb{N}$:

$$-\frac{\partial v_n}{\partial t} - \mathcal{L}v_n - f(\cdot, v_n, \sigma^\intercal D_x v_n) - n\int_E h^-(\mathcal{H}^e v_n - v_n, e)\lambda(de) \;\;=\;\; 0, \qquad (4.2)$$

on $[0, T) \times \mathbb{R}^d$.

To complete the PDE characterization of the function v, we need to provide a suitable boundary condition. In general, we can not expect to have $v(T^-, .) = g$, and we shall consider the relaxed boundary condition given by the equation:

$$\min\Big[v(T^-, \cdot) - g\;,\; \inf_{e\in E} h(\mathcal{H}^e v(T^-, \cdot) - v(T^-, \cdot), e)\Big] \;\;=\;\; 0, \quad \text{on } \mathbb{R}^d, \qquad (4.3)$$

In the sequel, we shall assume in addition to the conditions of paragraph 2.1 that the functions γ, f, c, and h are continuous with respect to all their arguments.

4.1 Viscosity properties

Solutions of (4.1), (4.2) and (4.3) are considered in the (discontinuous) viscosity sense, and it will be convenient in the sequel to define the notion of viscosity solutions in terms of sub- and super-jets. For a locally bounded function u on $[0, T] \times \mathbb{R}^d$, we define its lower semicontinuous (lsc in short) u_*, and upper semicontinuous (usc in short) envelope u^* by

$$u_*(t, x) \;=\; \liminf_{(t',x')\to(t,x),t'<T} u(t', x'), \qquad u^*(t, x) \;=\; \limsup_{(t',x')\to(t,x),t'<T} u(t', x').$$

Definition 4.1 *(Subjets and superjets)*
(i) For a function $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, lsc (resp. usc), we denote by $J^-u(t, x)$ the parabolic subjet (resp. $J^+u(t, x)$ the parabolic superjet) of u at $(t, x) \in [0, T] \times \mathbb{R}^d$, as the set of triples $(p, q, M) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d$ satisfying

$$\begin{aligned} u(t', x') \;\;\geq\;\; (\textit{resp.} \leq) \quad & u(t, x) + p(t' - t) + \langle q, x' - x\rangle + \frac{1}{2}\langle x' - x, M(x' - x)\rangle \\ & + o(|t' - t| + |x' - x|^2). \end{aligned}$$

(ii) For a function $u : [0, T) \times \mathbb{R}^d \to \mathbb{R}$, lsc (resp. usc), we denote by $\bar{J}^-u(t, x)$ the parabolic limiting subjet (resp. $\bar{J}^+u(t, x)$ the parabolic limiting superjet) of u at $(t, x) \in [0, T] \times \mathbb{R}^d$, as the set of triples $(p, q, M) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d$ such that

$$\begin{aligned} (p, q, M) \;&=\; \lim_n (p_n, q_n, M_n), \qquad (t, x) \;=\; \lim_n (t_n, x_n), \\ \textit{with} \quad (p_n, q_n, M_n) \in J^-u(t_n, x_n)\; (\textit{resp. } J^+u(t_n, x_n)), \qquad u(t, x) \;&=\; \lim_n u(t_n, x_n). \end{aligned}$$

We now give the definition of viscosity solutions to (4.1), (4.2) and (4.3).

Definition 4.2 *(Viscosity solutions to* (4.1)*)*
(i) A function u, lsc (resp. usc) on $[0,T) \times \mathbb{R}^d$*, is called a viscosity supersolution (resp. subsolution) to* (4.1) *if for each* $(t,x) \in [0,T) \times \mathbb{R}^d$*, and any* $(p,q,M) \in \bar{J}^- u(t,x)$ *(resp.* $\bar{J}^+ u(t,x)$*), we have*

$$\min\Big[-p - \langle b(x), q\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x)M) - f(x,u(t,x),\sigma^\intercal(x)q)\ , \\ \inf_{e\in E} h(\mathcal{H}^e[t,x,q,u] - u(t,x), e)\Big] \quad \geq (\text{ resp. } \leq) \quad 0.$$

(ii) A locally bounded function on $[0,T) \times \mathbb{R}^d$ *is called a viscosity solution to* (4.1) *if* u_* *is a viscosity supersolution and* u^* *is a viscosity subsolution to* (4.1).

Definition 4.3 *(Viscosity solutions to* (4.2)*)*
(i) A function u, lsc (resp. usc) on $[0,T) \times \mathbb{R}^d$*, is called a viscosity supersolution (resp. subsolution) to* (4.2) *if for each* $(t,x) \in [0,T) \times \mathbb{R}^d$*, and any* $(p,q,M) \in \bar{J}^- u(t,x)$ *(resp.* $\bar{J}^+ u(t,x)$*), we have*

$$-p - \langle b(x), q\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x)M) - f(x,u(t,x),\sigma^\intercal(x)q) \\ -n\int_E h^-(\mathcal{H}^e[t,x,q,u] - u(t,x), e)\lambda(de) \quad \geq (\text{ resp. } \leq) \quad 0.$$

(ii) A locally bounded function u on $[0,T) \times \mathbb{R}^d$ *is called a viscosity solution to* (4.2) *if* u_* *is a viscosity supersolution and* u^* *is a viscosity subsolution to* (4.2).

Definition 4.4 *(Viscosity solutions to* (4.3)*)*
(i) A function u, lsc (resp. usc) on $[0,T] \times \mathbb{R}^d$*, is called a viscosity supersolution (resp. subsolution) to* (4.3) *if for each (*$x \in \mathbb{R}^d$*, and any* $(p,q,M) \in \bar{J}^- u(T,x)$ *(resp.* $\bar{J}^+ u(T,x)$*), we have*

$$\min\Big[u(T,x) - g(x),\ \inf_{e\in E} h(\mathcal{H}^e[T,x,q,u] - u(T,x), e)\Big] \quad \geq (\text{ resp. } \leq) \quad 0.$$

(ii) A locally bounded function u on $[0,T] \times \mathbb{R}^d$ *is called a viscosity solution to* (4.3) *if* u_* *is a viscosity supersolution and* u^* *is a viscosity subsolution to* (4.3).

Remark 4.1 An equivalent definition of viscosity super and subsolution to (4.3), which shall be used later, is the following in terms of test functions : a function u, lsc (resp. usc) on $[0,T] \times \mathbb{R}^d$, is called a viscosity supersolution (resp. subsolution) to (4.3) if for each (t,x) $\in [0,T) \times \mathbb{R}^d$, and any $\varphi \in C^{1,2}([0,T] \times \mathbb{R}^d)$ such that (t,x) is a minimum (resp. maximum) global of $u - \varphi$, we have

$$\min\Big[u(T,x) - g(x),\ \inf_{e\in E} h(\mathcal{H}^e[T,x,D_x\varphi(T,x),u] - u(T,x), e)\Big] \quad \geq (\text{ resp. } \leq) \quad 0.$$

We have similar equivalent definitions of viscosity super and subsolution to (4.1) in terms of test functions.

From the Markov property of the jump-diffusion process X, and uniqueness of a minimal solution Y to the BSDE (2.10), we see that $Y_t = v(t, X_t)$, $0 \leq t \leq T$, where

$$v(t,x) \quad := \quad Y_t^{t,x}, \quad (t,x) \in [0,T] \times \mathbb{R}^d, \tag{4.4}$$

is a deterministic function of (t,x), $\{X_s^{t,x}, t \leq s \leq T\}$ is the solution to (2.1) starting from x at time t, and $\{Y_s^{t,x}, t \leq s \leq T\}$ is the minimal solution to (2.10)-(2.11) with $X_s = X_s^{t,x}$, $t \leq s \leq T$. Similarly, we define the function

$$v_n(t,x) \quad := \quad Y_t^{n,t,x}, \quad (t,x) \in [0,T] \times \mathbb{R}^d, \tag{4.5}$$

where $\{(Y_s^{n,t,x}, Z_s^{n,t,x}, U_s^{n,t,x}(.)), t \leq s \leq T\}$ is the unique solution to (3.1) with $X_s = X_s^{t,x}$, $t \leq s \leq T$. The relation between the penalized BSDE (3.1) and the penalized IPDE (4.2) is well-known from the results of [2]. Although our framework does not fit exactly into the one of [2], by mimicking closely the arguments in this paper and using comparison theorem in [20], we obtain the following result.

Proposition 4.1 *The function v_n in* (4.5) *is a continuous viscosity solution to* (3.1).

We slightly strengthen Assumption **(H1)** or **(H2)** by

(H1') There exists a quadruple $(\tilde{Y}, \tilde{Z}, \tilde{K}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{A^2}$ satisfying (2.12), with $\tilde{Y}_t = \tilde{v}(t, X_t)$, $0 \leq t \leq T$, for some function deterministic $\tilde{v}$ satisfying a linear growth condition

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|\tilde{v}(t,x)|}{1 + |x|} < +\infty$$

(H2') There exists a quadruple $(\tilde{Y}, \tilde{Z}, \tilde{K}, \tilde{U}) \in \mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ satisfying (2.10)-(2.11), with $\tilde{Y}_t = \tilde{v}(t, X_t)$, $0 \leq t \leq T$, for some function deterministic $\tilde{v}$ satisfying a linear growth condition

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|\tilde{v}(t,x)|}{1 + |x|} < +\infty$$

Remark 4.2 Assumption **(H2')** (or **(H1')** which is weaker than **(H2')** in the case $h(u,e) = -u$) ensures that the function v in (4.4) satisfies a linear growth condition, and is in particular locally bounded. Indeed, from (3.10) and by passing to the limit by Fatou's lemma for $v(t,x) = Y_t^{t,x} = \lim Y_t^{n,t,x}$, we have

$$\sup_{t \in [0,T]} |v(t,x)|^2 \quad \leq \quad C\Big(1 + \mathbf{E}|g(X_T^{t,x})|^2 + \mathbf{E}\Big[\int_t^T |X_s^{t,x}|^2 dt\Big] + \mathbf{E}\big[\sup_{s \in [t,T]} |\tilde{v}(s, X_s^{t,x}|^2\big]\Big).$$

The result follows from the standard estimate

$$\mathbf{E}[\sup_{t \leq s \leq T} |X_s^{t,x}|^2] \quad \leq \quad C(1 + |x|^2),$$

and the linear growth conditions on g and $\tilde{v}$.

By adapting stability arguments for viscosity solutions to our context, we now prove the viscosity property of the function v to (4.1).

Theorem 4.1 *Under* **(H2')** *(or* **(H1')** *in the case : $h(u,e) = -u$), the function v in* (4.4) *is a viscosity solution to* (4.1).

Proof. From the results of the previous section, we know that v is the pointwise limit of the nondecreasing sequence of functions (v_n). By continuity of v_n, we then have (see e.g. [1] p. 91) :

$$v \;=\; v_* \quad = \quad \lim_{n\to\infty} \inf{}_* v_n, \quad \text{where} \quad \lim_{n\to\infty} \inf{}_* v_n(t,x) \;:=\; \liminf_{\substack{n\,\to\,\infty \\ t'\,\to\,t,\,x'\,\to\,x}} \; v_n(t',x'), \quad (4.6)$$

$$v^* \quad = \quad \lim_{n\to\infty} \sup{}^* v_n, \quad \text{where} \quad \lim_{n\to\infty} \sup{}^* v_n(t,x) \;:=\; \limsup_{\substack{n\,\to\,\infty \\ t'\,\to\,t,\,x'\,\to\,x}} \; v_n(t',x'). \quad (4.7)$$

(i) We first show the viscosity supersolution property for $v = v_*$. Let (t,x) a point in $[0,T) \times \mathbb{R}^d$, and $(p,q,M) \in \bar{J}^- v(t,x)$. By (4.6) and Lemma 6.1 in [7], there exists sequences

$$n_j \;\to\; \infty, \quad (p_j, q_j, M_j) \;\in\; J^- v_{n_j}(t_j, x_j),$$

such that

$$(t_j, x_j, v_{n_j}(t_j,x_j), p_j, q_j, M_j) \quad \to \quad (t, x, v(t,x), p, q, M). \tag{4.8}$$

We also have by definition of $v = v_*$ and continuity of γ :

$$v(t, x+\gamma(x,e)) \quad \leq \quad \liminf_{j\to\infty} v_{n_j}(t_j, x_j + \gamma(x_j,e)), \quad \forall e \in E. \tag{4.9}$$

Moreover, from the viscosity supersolution property for v_{n_j}, we have for all j

$$\begin{aligned} & - p_j - \langle b(x_j), q_j\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x_j)M_j) - f(x_j, v_{n_j}(t_j,x_j), \sigma^{\intercal}(x_j)q_j) \\ & \qquad -n_j \int_E h^-(\mathcal{H}^e[t_j, x_j, q_j, v_{n_j}] - v_{n_j}(t_j,x_j), e)\lambda(de) \quad \geq \quad 0. \end{aligned} \tag{4.10}$$

Let us check that the following inequality holds :

$$\inf_{e\in E} h(\mathcal{H}^e[t,x,q,v] - v(t,x), e) \quad \geq \quad 0. \tag{4.11}$$

We argue by contradiction, and assume there exists some $e_0 \in E$ s.t.

$$h(v(t, x+\gamma(x,e_0))) + c(x, v(t,x), \sigma^{\intercal}(x)q, e_0) - v(t,x), e_0) \quad < \quad 0.$$

Then, by continuity of σ, h, γ, c in all their variables, (4.8), (4.9), and the nonincreasing property of h, one may find some $\varepsilon > 0$ and some open neighborhood $\mathcal{O}_0$ of e_0 such that for all j large enough :

$$h(v_{n_j}(t_j, x_j + \gamma(x_j,e)) + c(x_j, v_{n_j}(t_j,x_j), \sigma^{\intercal}(x_j)q_j, e) - v_{n_j}(t_j,x_j), e) \quad \leq \quad -\varepsilon, \quad \forall e \in \mathcal{O}_0.$$

Since the support of λ is E, this implies

$$\int_E h^-(\mathcal{H}^e(t_j, x_j, q_j, v_{n_j}) - v_{n_j}(t_j,x_j), e)\lambda(de) \quad \geq \quad \varepsilon\lambda(\mathcal{O}_0) \;>\; 0.$$

By sending j to infinity into (4.10), we get the required contradiction. On the other hand, by (4.10), we have

$$-p_j - \langle b(x_j), q_j\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x_j)M_j) - f(x_j, v_{n_j}(t_j,x_j), \sigma^{\intercal}(x_j)q_j) \quad \geq \quad 0,$$

so that by sending j to infinity:

$$-p - \langle b(x), q\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x)M) - f(x, v(t,x), \sigma^{\intercal}(x)q) \quad \geq \quad 0,$$

which proves, together with (4.11), that v is a viscosity supersolution to (4.1).
(ii) We conclude by showing the viscosity subsolution property for v^*. Let (t,x) a point in $[0,T)\times\mathbb{R}^d$, and $(p,q,M)\in\bar{J}^+v^*(t,x)$ such that

$$\inf_{e\in E} h(\mathcal{H}^e[t,x,q,v^*] - v^*(t,x), e) \quad > \quad 0. \tag{4.12}$$

From (4.7) and Lemma 6.1 in [7], there exists sequences

$$n_j \;\to\; \infty, \quad (p_j,q_j,M_j) \;\in\; J^+v_{n_j}(t_j,x_j),$$

such that

$$(t_j,x_j,v_{n_j}(t_j,x_j),p_j,q_j,M_j) \quad\to\quad (t,x,v^*(t,x),p,q,M). \tag{4.13}$$

By continuity of the functions c,γ, and definition of v^*, we also have

$$\limsup_{j\to\infty} \mathcal{H}^e[t_j,x_j,q_j,v_{n_j}] \quad\leq\quad \mathcal{H}^e[t,x,q,v^*], \quad \forall\, e\in E. \tag{4.14}$$

Now, from the viscosity subsolution property for v_{n_j}, we have for all j

$$\begin{aligned} &- p_j - \langle b(x_j), q_j\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x_j)M_j) - f(x_j, v_{n_j}(t_j,x_j), \sigma^{\intercal}(x_j)q_j) \\ &\qquad -n_j\int_E h^-(\mathcal{H}^e[t_j,x_j,q_j,v_{n_j}] - v_{n_j}(t_j,x_j), e)\lambda(de) \quad\leq\quad 0. \end{aligned} \tag{4.15}$$

From (4.12)-(4.13)-(4.14), continuity assumptions on h,c, and the nonincreasing property of h, we have for j large enough

$$h(\mathcal{H}^e[t_j,x_j,q_j,v_{n_j}] - v_{n_j}(t_j,x_j), e) \quad > \quad 0, \quad \forall e\in E,$$

and so

$$\int_E h^-(\mathcal{H}^e[t_j,x_j,q_j,v_{n_j}] - v_{n_j}(t_j,x_j), e)\lambda(de) \quad = \quad 0.$$

Hence, by taking the limit as j goes to infinity, into (4.15), we conclude that

$$-p - \langle b(x), q\rangle - \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x)M) - f(x, v^*(t,x), \sigma^{\intercal}(x)q) \quad\leq\quad 0,$$

which shows the viscosity subsolution property for v^* to (4.1). □

We next turn to the boundary condition.

Theorem 4.2 *Under* **(H2')** *(or* **(H1')** *in the case : $h(u,e) = -u$), the function v in* (4.4) *is a viscosity solution to* (4.3).

In order to deal with the possible jump at the terminal condition, we need the following dynamic programming caracterization of the minimal solution.

Lemma 4.1 *Let $(t,x) \in [0,T) \times \mathbb{R}^d$, and $(Y^{t,x}, Z^{t,x}, U^{t,x}, K^{t,x})$ be a minimal solution to* (2.10)-(2.11) *on $[t,T]$ with $X_s = X_s^{t,x}$. Then for any stopping time θ valued in $[t,T]$, $(Y_s^{t,x}, Z_s^{t,x}, U_s^{t,x}, K_s^{t,x})_{s\in[t,\theta]}$ is a minimal solution to :*

$$\begin{aligned} Y_s &= v(\theta, X_\theta^{t,x}) + \int_s^\theta f(X_r^{t,x}, Y_r, Z_r)dr + K_\theta^{t,x} - K_s^{t,x} \qquad (4.16) \\ &\quad - \int_s^\theta \langle Z_r, dW_r\rangle - \int_s^\theta \int_E \Big(U_r(e) - c(X_{r^-}^{t,x}, Y_{r^-}, Z_r, e)\Big)\,\mu(dr,de) \end{aligned}$$

with

$$h(U_s(e),e) \geq 0 \quad d\mathbf{P} \otimes dt \otimes \lambda(de) \quad a.e.\ on\ \Omega \times [t,\theta] \times E. \qquad (4.17)$$

Proof. Let Y^1 be the minimal solution on $[t,\theta]$ of (4.16)-(4.17) (the existence of a minimal solution in the case of a random terminal time is obtained by similar arguments to those used in the case of a deterministic terminal time). For each $\omega \in \Omega$, there exists a minimal solution $Y^{2,\omega}$ on $[\theta(\omega), T]$ to (2.10)-(2.11). We then have from the Markov property of X that $Y^{2,\omega}_{\theta(\omega)} = v\left(\theta(\omega), X^{t,x}_{\theta(\omega)}(\omega)\right)$ for all $\omega \in \Omega$. By a measurable selection theorem, there exists $Y^2 \in \mathcal{S}^{\mathbf{2}}$ such that for all $\omega \in \Omega$ we have $Y^2_{\theta(\omega)}(\omega) = Y^{2,\omega}_{\theta(\omega)} = v\left(\theta(\omega), X^{t,x}_{\theta(\omega)}(\omega)\right)$ and $Y^2_s(\omega) = Y^{2,\omega}_s(\omega)$ for $s \in [\theta(\omega), T]$. We then define the process $\tilde{Y}$ by $\tilde{Y}|_{[t,\theta]} = Y^1$ and $\tilde{Y}|_{(\theta,T]} = Y^2$. Hence, $\tilde{Y}$ is a solution on $[t,T]$ to (2.10)-(2.11), which implies $\tilde{Y} \geq Y^{t,x}$. Moreover, since $Y^{t,x}_\theta = v(\theta, X^{t,x}_\theta)$, it follows that $(Y_s^{t,x}, Z_s^{t,x}, U_s^{t,x}, K_s^{t,x})_{s\in[t,\theta]}$ is a solution on $[t,\theta]$ to (4.16)-(4.17). Hence $Y^1 \leq Y^{t,x}$ on $[t,\theta]$, and therefore $Y^1 = Y^{t,x}$ on $[t,\theta]$. □

Proof of Theorem 4.2 (i) We first prove the supersolution property of v_* to (4.3). Let $x \in \mathbb{R}^d$, and $(p,q,M) \in \bar{J}^- v_*(T,x)$. By same arguments as in (4.11), we have

$$\inf_{e\in E} h(\mathcal{H}^e[T,x,q,v_*] - v_*(T,x), e) \quad \geq \quad 0. \qquad (4.18)$$

Moreover, since the sequence of continuous functions $(v_n)_n$ in nondecreasing and $v_n(T,.) = g$, we deduce that $v_*(T,.) \geq g$, which combined with (4.18), proves the viscosity supersolution property for v_* to (4.3).

(ii) We next prove the subsolution property of v^* to (4.3). We argue by contradiction and asume that there exist $x_0 \in \mathbb{R}^n$, $\varphi \in C^{1,2}([0,T] \times \mathbb{R}^n)$ such that

$$0 \ = \ (v^* - \varphi)(T, x_0) \quad = \quad \max_{[0,T]\times\mathbb{R}^d} (v^* - \varphi) \qquad (4.19)$$

and

$$\min\left[\varphi(T,x_0) - g(x_0)\ ,\ \inf_{e\in E} h(\mathcal{H}^e[T,x_0,D_x\varphi(T,x_0),v^*] - \varphi(T,x_0), e)\right] \quad =: \quad 2\varepsilon \ > \ 0.$$

By the upper semicontinuity of v^*, the continuity of φ and its derivative, and the nonincreasing property of h, there exists an open neighbohood $\mathcal{O}$ of (T, x_0) in $[0,T] \times \mathbb{R}^d$, and $A, r > 0$ such that for all $(t,x,\alpha,\beta) \in \mathcal{O} \times (-A, A) \times B(0,r)$, we have

$$\begin{aligned} \varepsilon \quad \leq \quad & \min \Big[\varphi(t,x) - \alpha - g(x) \ , \\ & \inf_{e \in E} h(v^*(t, x+\gamma(x,e))+c(x,\varphi(t,x)-\alpha, \sigma^\intercal(x)[D_x\varphi(t,x)+\beta]) - [\varphi(t,x)-\alpha], e) \Big]. \end{aligned} \tag{4.20}$$

Let $(t_k, x_k)_k$ be a sequence in $[0,T) \times \mathbb{R}^d$ such that

$$(t_k, x_k) \to (T, x_0) \quad \text{ and } \quad v(t_k, x_k) \to v^*(T, x_0). \tag{4.21}$$

Fix then $\delta > 0$ such that for k large enough: $[t_k, T] \times B(x_k, \delta) \subset \mathcal{O}$, and let us define the functions φ_k by

$$\varphi_k(t,x) \quad = \quad \varphi(t,x) + \zeta \frac{|x-x_k|^2}{\delta^2} + C_k \phi\left(\frac{x-x_k}{\delta}\right) + \sqrt{T-t},$$

where $0 < \zeta < A \wedge \delta r$, $\phi \in C^2(\mathbb{R}^d)$ satisfies $\phi|_{\bar{B}(0,1)} \equiv 0$, $\phi|_{\bar{B}(0,1)^c} > 0$ and $\lim_{|x| \to \infty} \frac{\phi(x)}{1+|x|} = \infty$, and $C_k > 0$ is a constant to be chosen below. By (4.19), we notice that

$$(v^* - \varphi_k)(t,x) \quad \leq \quad -\zeta \qquad \text{for } \ (t,x) \in [t_k, T] \times \partial B(x_k, \delta),$$

and from the conditions on ϕ, we can choose C_k (large enough) so that

$$(v^* - \varphi_k)(t,x) \quad \leq \quad -\frac{\zeta}{2} \qquad \text{for } (t,x) \in B(x_k, \delta)^c \times [t_k, T]. \tag{4.22}$$

Since $\frac{\partial}{\partial t}(\sqrt{T-t}) \to -\infty$ as $t \nearrow T$, we have for k large enough :

$$\begin{aligned} & -\frac{\partial \varphi_k}{\partial t} - \mathcal{L}\varphi_k(t,x) - f(x, \varphi_k(t,x) - \alpha, \sigma^\intercal(x) D_x \varphi_k(t,x)) \\ \geq & \quad 0 \quad \text{ for } \ (t,x,\alpha) \in [t_k, T) \times B(x_k, \delta) \times (-A+\zeta, A). \end{aligned} \tag{4.23}$$

Fix now $\alpha^* \in (0, A \wedge \frac{\zeta}{2} \wedge \varepsilon)$, and let us denote $\tau_k = \inf\big\{s \geq t_k \ ; \ X_s^k \neq X_{s^-}^k\big\}$, $\theta_k = \inf\big\{s \geq t_k \ ; \ X_s^k \notin B(x_k, \delta)\big\} \wedge \tau_k \wedge T$ where $X^k = X^{t_k, x_k}$. Let us then define the quadruples (Y^k, Z^k, U^k, K^k) on $[t_k, \theta_k]$ by :

$$Y_s^k = \Big[\varphi_k(s, X_s^k) - \alpha^*\Big] \mathbf{1}_{\{s \in [t_k, \theta_k)\}} + v(\theta_k, X_{\theta_k}^k) \mathbf{1}_{\{s = \theta_k\}} \ , \ Z_s^k = \sigma^\intercal(X_{s^-}^k) D_x \varphi_k(s, X_{s^-}^k) \ ,$$

$$\begin{aligned} U_s^k(e) \quad = \quad & v^*(s, X_{s^-}^k + \gamma(X_{s^-}^k, e)) + c(X_{s^-}^k, \varphi_k(s, X_{s^-}^k) - \alpha^*, \sigma^\intercal(X_{s^-}^k) D_x \varphi_k(s, X_{s^-}^k)) \\ & - [\varphi_k(s, X_{s^-}^k) - \alpha^*], \end{aligned}$$

and

$$\begin{aligned} K_s^k \quad = \quad & - \int_{t_k}^s \left\{ \frac{\partial \varphi_k}{\partial t}(r, X_r^k) + \mathcal{L}\varphi_k(r, X_r^k) + f(X_r^k, \varphi_k(r, X_r^k) - \alpha^*, \sigma^\intercal(X_r^k) D_x \varphi_k(r, X_r^k)) \right\} dr \\ & - \int_{t_k}^s \int_E (\varphi_k - \alpha^* - v^*)(r, X_{r^-}^k + \gamma(X_{r^-}^k, e)) \mu(dr, de) \\ & + \Big(\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* - v(\theta_k, X_{\theta_k}^k)\Big) \mathbf{1}_{\{s = \theta_k\}}. \end{aligned}$$

By construction and from Itô's formula on $\varphi_k(s, X_s^k)$, we see that (Y^k, Z^k, U^k, K^k) satisfies (4.16) on $[t_k, \theta_k]$. From (4.20), it is clear that the process U^k satisfies the constraint :

$$h(U_t^k(e), e) \quad \geq \quad 0, \quad d\mathbf{P} \otimes dt \otimes \lambda(de) \;\; a.e. \;\; \text{on } \Omega \times [t_k, \theta_k] \times E.$$

Observe also that

$$\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* \;\geq\; v(\theta_k, X_{\theta_k}^k) \tag{4.24}$$

Indeed, we have two cases:

- $\theta_k < T$: in this case $(\theta_k, X_{\theta_k}^k) \notin \mathcal{O}$, and since $\alpha^* < \frac{\zeta}{2}$, we have by (4.22),

$$\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* \;\geq\; v^*(\theta_k, X_{\theta_k}^k) \;\geq\; v(\theta_k, X_{\theta_k}^k).$$

- $\theta_k = T$: in this case $(\theta_k, X_{\theta_k}^k) = (T, X_T^k) \in \mathcal{O}$. Since $\alpha^* \leq \varepsilon$, we have by (4.20)

$$\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* \;\geq\; \varphi(\theta_k, X_{\theta_k}^k) - \varepsilon \geq g(X_T^k) \;=\; v(\theta_k, X_{\theta_k}^k).$$

Let us then check that K^k is nondecreasing on $[t_k, \theta_k]$. First, on $[t_k, \theta_k)$, we notice that K^k consists only in the Lebesgue term dr, and so is nondecreasing by (4.23). Moreover, we see that $K_{\theta_k}^k \geq K_{\theta_k^-}^k$. Indeed, there are two possible cases:

- $\theta_k < \tau_k$: then $K_{\theta_k}^k = K_{\theta_k^-}^k + \varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* - v(\theta_k, X_{\theta_k}^k)$, and by (4.24), we have $K_{\theta_k}^k \geq K_{\theta_k^-}^k$.

- $\theta_k = \tau_k$: then $K_{\theta_k}^k = K_{\theta_k^-}^k - (\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* - v^*(\theta_k, X_{\theta_k}^k)) + (\varphi_k(\theta_k, X_{\theta_k}^k) - \alpha^* - v(\theta_k, X_{\theta_k}^k))$, and so $K_{\theta_k}^k \geq K_{\theta_k^-}^k$.

Therefore, the quadruple (Y^k, Z^k, U^k, K^k) is a solution on $[t_k, \theta_k]$ to (4.16)-(4.17), and by Lemma 4.1, we deduce that for all k,

$$\varphi_k(t_k, x_k) - \alpha^* \;=\; \varphi(t_k, x_k) + \sqrt{T - t_k} - \alpha^* \;\geq\; v(t_k, x_k).$$

We finally obtain a contradiction by sending k to ∞. $\Box$

4.2 Uniqueness result

This paragraph is devoted to a uniqueness result for the QVI (4.1)-(4.3). We need to impose some additional assumptions.

(H3) There exists a nonnegative function $\Lambda \in \mathcal{C}^2(\mathbb{R}^d)$ and a positive constant ρ satisfying

(i) $\mathcal{L}\Lambda + f(., \Lambda, \sigma^\intercal D\Lambda) \leq \rho\Lambda$,

(ii) $\inf_{e \in E} h(\mathcal{H}^e\Lambda(x) - \Lambda(x), e) > 0$ for all $x \in \mathbb{R}^d$,

(iii) $\Lambda(x) \geq g(x)$ for all $x \in \mathbb{R}^d$,

(iv) $\lim_{|x| \to \infty} \frac{\Lambda(x)}{1+|x|} = \infty$.

Assumption **(H3)** essentially ensures the existence of a suitable strict supersolution to (4.1). We shall give in paragraph 6 some sufficient conditions for **(H3)**. This strict supersolution allows to control the nonlocal term in QVI (4.1)-(4.3) via some convex small perturbation. Thus, to deal with the dependence of f, c on y, z, we also require some convexity conditions.

(H4)

(i) The function $f(x,.,.)$ is convex in $(y,z) \in \mathbb{R} \times \mathbb{R}^d$ for all $x \in \mathbb{R}^d$.

(ii) The function $h(.,e)$ is concave in $u \in \mathbb{R}$ a for all $e \in E$.

(iii) The function $c(x,.,.,e)$ is convex in $(y,z) \in \mathbb{R} \times \mathbb{R}^d$ for all $(x,e) \in \mathbb{R}^d \times E$.

(iv) The function $c(x,.,z,e)$ is decreasing in $y \in \mathbb{R}$ for all $(x,z,e) \in \mathbb{R}^d \times \mathbb{R}^d \times E$.

Theorem 4.3 *Assume that* **(H3)** *and* **(H4)** *hold, and let U (resp. V) be a lsc (resp. usc) viscosity supersolution (resp. subsolution) to* (4.1)-(4.3) *satisfying a linear growth condition :*

$$\sup_{x\in\mathbb{R}^d} \frac{|U(t,x)| + |V(t,x)|}{1+|x|} \quad < \quad \infty, \quad \forall t \in [0,T].$$

Then, $U \geq V$ on $[0,T] \times \mathbb{R}^d$. Consequently, under **(H2')** *(or* **(H1')** *in the case : $h(u,e) = -u$),* **(H3)** *and* **(H4)**, *the function v in* (4.4) *is the unique viscosity solution to* (4.1)-(4.3) *satisfying a linear growth condition, and v is continuous on $[0,T) \times \mathbb{R}^d$.*

Proof. • *Comparison principle.* As usual, we shall argue by contradiction by assuming that

$$\sup_{[0,T]\times\mathbb{R}^d} (V - U) > 0. \tag{4.25}$$

1. For some $\lambda > 0$ to be chosen below, let

$$\tilde{U}(t,x) = e^{(\rho+\lambda)t}U(t,x)\ ,\ \tilde{V}(t,x) = e^{(\rho+\lambda)t}V(t,x) \ \text{ and } \ \tilde{\Lambda}(t,x) = e^{(\rho+\lambda)t}\Lambda(x).$$

A straightforward derivation shows that $\tilde{U}$ (resp. $\tilde{V}$) is a viscosity supersolution (resp. subsolution) to

$$\begin{aligned} \min\Big[\rho w - \frac{\partial w}{\partial t} - \mathcal{L}w - \tilde{f}\left(\cdot, w, \sigma^{\intercal} D_x w\right), & \\ \inf_{e\in E} \tilde{h}\left(\cdot, \tilde{\mathcal{H}}^e w - w, e\right)\Big] &= 0, \ \text{ on } [0,T) \times \mathbb{R}^d \end{aligned} \tag{4.26}$$

$$\min\Big[w(T^-,\cdot) - \tilde{g}\ ,\ \inf_{e\in E} \tilde{h}(T, \tilde{\mathcal{H}}^e w(T^-,\cdot) - w(T^-,\cdot), e)\Big] \;=\; 0 \ \text{ on } \mathbb{R}^d \tag{4.27}$$

where

$$\begin{aligned} \tilde{f}(t,x,r,q) &= e^{(\rho+\lambda)t} f\left(x, re^{-(\rho+\lambda)t}, qe^{-(\rho+\lambda)t}\right) - \lambda r \\ \tilde{h}(t,r,e) &= e^{(\rho+\lambda)t} h(e^{-(\rho+\lambda)t} r, e), \quad \tilde{g}(x) \;=\; e^{(\rho+\lambda)T} g(x) \end{aligned}$$

and

$$\tilde{\mathcal{H}}w(t,x) \quad = \quad w(t,x+\gamma(x,e)) + \tilde{c}\left(x, w(t,x), \sigma^{\intercal}(x)D_x w(t,x), e\right)$$

with

$$\tilde{c}(t,x,r,q,e) \quad = \quad e^{(\rho+\lambda)t}c(x, e^{-(\rho+\lambda)t}r, e^{-(\rho+\lambda)t}q, e)$$

for all $(t,x,r,q,e) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times E$. Since f is Lipschitz, we can choose λ large enough so that $\tilde{f}$ is nonincreasing in r. Denote $\tilde{W} = (1-\mu)\tilde{U} + \mu\tilde{\Lambda}$ with $\mu > 0$. By (4.25) and the growth condition **(H3)**(iv) of Λ, we have for μ small enough

$$\sup_{[0,T]\times\mathbb{R}^d}(\tilde{V}-\tilde{W}) = (\tilde{V}-\tilde{W})(t_0,x_0) > 0. \tag{4.28}$$

for some $(t_0,x_0) \in [0,T]\times\mathbb{R}^d$. Moreover from the viscosity supersolution property (4.26)-(4.27) of $\tilde{U}$, and the conditions **(H3)**(i), (ii), **(H4)**(i), (ii), (iii), we see that $\tilde{W}$ is a viscosity supersolution to

$$\rho w - \frac{\partial w}{\partial t} - \mathcal{L}w - \tilde{f}\left(\cdot, w, \sigma^{\intercal}D_x w\right) \quad \geq \quad 0, \quad \text{ on } [0,T)\times\mathbb{R}^d, \tag{4.29}$$

$$\inf_{e\in E}\tilde{h}\left(\cdot, \tilde{\mathcal{H}}^e w - w, e\right) \quad \geq \quad \mu\tilde{q}, \quad \text{ on } [0,T]\times\mathbb{R}^d, \tag{4.30}$$

where $\tilde{q}(t,x) = e^{(\rho+\lambda)t}\inf_{e\in E}h(\mathcal{H}^e\Lambda(x)-\Lambda(x),e)$ is positive on $[0,T]\times\mathbb{R}^d$ by **(H3)**(ii).

2. Denote for all $(t,x,y)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ and $n\geq 1$

$$\Theta_n(t,x,y) \quad = \quad \tilde{V}(t,x) - \tilde{W}(t,y) - \varphi_n(t,x,y),$$

with

$$\varphi_n(t,x,y) \;=\; n|x-y|^2 + |x-x_0|^4 + |t-t_0|^2.$$

By the growth assumption on U and V and **(H3)**(iii), for all n, there exists $(t_n,x_n,y_n)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ attaining the maximum of Θ_n on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$. By standard arguments, we have :

$$(t_n,x_n,y_n) \quad \to \quad (t_0,x_0,x_0), \tag{4.31}$$

$$n|x_n-y_n|^2 \quad \to \quad 0, \tag{4.32}$$

$$\tilde{V}(t_n,x_n)-\tilde{W}(t_n,y_n) \quad \to \quad \tilde{V}(t_0,x_0)-\tilde{W}(t_0,x_0). \tag{4.33}$$

3. We now show that for n large enough

$$\inf_{e\in E}\tilde{h}(t_n, \tilde{\mathcal{H}}^e[t_n,x_n,D_x\varphi_n(t_n,x_n,y_n),\tilde{V}] - \tilde{V}(t_n,x_n),e) \quad > \quad 0. \tag{4.34}$$

On the contrary, up to a subsequence, we would have for all n,

$$\inf_{e\in E}\tilde{h}(t_n, \tilde{\mathcal{H}}^e[t_n,x_n,D_x\varphi_n(t_n,x_n,y_n),\tilde{V}] - \tilde{V}(t_n,x_n),e) \;\leq\; 0,$$

and so by uppersemicontinuity of $\tilde{V}$, compactness of E, there would exist a sequence (e_n) in E such that

$$\tilde{h}(t_n, \tilde{\mathcal{H}}^{e_n}[t_n, x_n, D_x\varphi_n(t_n, x_n, y_n), \tilde{V}] - \tilde{V}(t_n, x_n), e_n) \quad \leq \quad 0.$$

Moreover, by the viscosity supersolution property of $\tilde{W}$ to (4.30), we have

$$\tilde{h}(t_n, \tilde{\mathcal{H}}^{e_n}[t_n, y_n, -D_y\varphi_n(t_n, x_n, y_n), \tilde{W}] - \tilde{W}(t_n, y_n), e_n) \quad \geq \quad \mu\tilde{q}(t_n, y_n).$$

From the nonincreasing and the Lipschitz property of $h(., e)$, we deduce from the two previous inequalities that there exists a positive constant η such that

$$\begin{aligned} & \tilde{\mathcal{H}}^{e_n}[t_n, y_n, -D_y\varphi_n(t_n, x_n, y_n), \tilde{W}] - \tilde{W}(t_n, y_n) + \eta\tilde{q}(t_n, y_n) \\ \leq \ & \tilde{\mathcal{H}}^{e_n}[t_n, x_n, D_x\varphi_n(t_n, x_n, y_n), \tilde{V}] - \tilde{V}(t_n, x_n), \end{aligned}$$

which is rewritten as

$$\begin{aligned} & \tilde{V}(t_n, x_n) - \tilde{W}(t_n, y_n) + \eta\tilde{q}(t_n, y_n) \\ \leq \ & \tilde{V}(t_n, x_n + \gamma(x_n, e_n)) - \tilde{W}(t_n, y_n + \gamma(y_n, e_n)) + \Delta C_n \end{aligned} \tag{4.35}$$

where

$$\begin{aligned} \Delta C_n \quad = \quad & \tilde{c}\left(t_n, x_n, \tilde{V}(t_n, x_n), \sigma^\intercal(x_n) D_x\varphi_n(t_n, x_n, y_n), e_n\right) \\ & - \tilde{c}\left(t_n, y_n, \tilde{W}(t_n, y_n), -\sigma^\intercal(y_n) D_y\varphi_n(t_n, x_n, y_n)\right). \end{aligned}$$

Now, we write $\Delta C_n = \Delta C_n^1 + \Delta C_n^2$, with

$$\begin{aligned} \Delta C_n^1 \quad = \quad & \tilde{c}\left(t_n, x_n, \tilde{V}(t_n, x_n), \sigma^\intercal(x_n) D_x\varphi_n(t_n, x_n, y_n), e_n\right) \\ & - \tilde{c}\left(t_n, x_n, \tilde{W}(t_n, y_n), \sigma^\intercal(x_n) D_x\varphi_n(t_n, x_n, y_n), e_n\right), \\ \Delta C_n^2 \quad = \quad & \tilde{c}\left(t_n, x_n, \tilde{W}(t_n, y_n), \sigma^\intercal(x_n) D_x\varphi_n(t_n, x_n, y_n), e_n\right) \\ & - \tilde{c}\left(t_n, x_n, \tilde{W}(t_n, y_n), -\sigma^\intercal(y_n) D_y\varphi_n(t_n, x_n, y_n), e_n\right), \\ \Delta C_n^3 \quad = \quad & \tilde{c}\left(t_n, x_n, \tilde{W}(t_n, y_n), -\sigma^\intercal(y_n) D_y\varphi_n(t_n, x_n, y_n), e_n\right) \\ & - \tilde{c}\left(t_n, y_n, \tilde{U}(t_n, y_n), -\sigma^\intercal(y_n) D_y\varphi_n(t_n, x_n, y_n), e_n\right). \end{aligned}$$

We have $\tilde{V}(t_n, x_n) - \tilde{W}(t_n, y_n) \to (\tilde{V} - \tilde{W})(t_0, x_0) > 0$ by (4.28) and (4.33). Hence, for n large enough, $\tilde{V}(t_n, x_n) \geq \tilde{W}(t_n, y_n)$, and so from the nonincreasing condition **(H4)**(iv) of c, we have $\Delta C_n^1 \leq 0$. Since $\sigma^\intercal(x_n) D_x\varphi_n(t_n, x_n, y_n) + \sigma^\intercal(y_n) D_y\varphi_n(t_n, x_n, y_n) \to 0$ by the Lipschitz condition on σ and (4.32), we deduce with the Lipschitz condition on c that $\limsup_{n\to\infty} \Delta C_n^2 \leq 0$. By (4.31) and continuity of c, we have $\lim_{n\to\infty} \Delta C_n^3 = 0$. Therefore, we obtain

$$\limsup_{n\to\infty} \Delta C_n \quad \leq \quad 0.$$

Up to a subsequence, we may assume that (e_n) converges to e_0 in E. Hence, by sending n to infinity into (4.35), it follows with (4.33) and the upper (resp. lower)-semicontinuity of $\tilde{V}$ (resp. $\tilde{W}$) that :

$$\begin{aligned}(\tilde{V}-\tilde{W})(t_0, x_0+\gamma(x_0,e_0), x_0+\gamma(x_0,e_0)) &\geq (\tilde{V}-\tilde{W})(t_0,x_0)+\eta\tilde{q}(t_0,x_0) \\ &> (\tilde{V}-\tilde{W})(t_0,x_0),\end{aligned}$$

a contradiction with (4.28).

4. Let us check that, up to a subsequence, $t_n < T$ for all n. On the contrary, $t_n = t_0 = T$ for n large enough, and from (4.34), and the viscosity subsolution property of $\tilde{V}$ to (4.27), we would get

$$\tilde{V}(T,x_n) \leq \tilde{g}(x_n).$$

On the other hand, by the viscosity supersolution property of $\tilde{U}$ to (4.27) and **(H3)**(iii), we have $\tilde{W}(T,y_n) \geq \tilde{g}(y_n)$, and so

$$\tilde{V}(T,x_n) - \tilde{W}(T,y_n) \leq \tilde{g}(x_n) - \tilde{g}(y_n).$$

By sending n to infinity, and from continuity of $\tilde{g}$, this would imply $(\tilde{V}-\tilde{W})(t_0,x_0) \leq 0$, a contradiction with (4.28).

5. We may then apply Ishii's lemma (see Theorem 8.3 in [7]) to $(t_n,x_n,y_n) \in [0,T)\times\mathbb{R}^d\times\mathbb{R}^d$ that attains the maximum of Θ_n, for all $n \geq 1$: there exist $(p^n_{\tilde{V}}, q^n_{\tilde{V}}, M_n) \in \bar{J}^{2,+}\tilde{V}(t_n,x_n)$ and $(p^n_{\tilde{W}}, q^n_{\tilde{W}}, N_n) \in \bar{J}^{2,-}\tilde{W}(t_n,y_n)$ such that

$$\begin{aligned} p^n_{\tilde{V}} - p^n_{\tilde{W}} &= \partial_t\varphi_n(t_n,x_n,y_n) = 2(t_n-t_0), \\ q^n_{\tilde{V}} &= D_x\varphi_n(t_n,x_n,y_n), \qquad q^n_{\tilde{W}} = -D_y\varphi_n(t_n,x_n,y_n),\end{aligned}$$

and

$$\begin{pmatrix} M_n & 0 \\ 0 & -N_n \end{pmatrix} \leq A_n + \frac{1}{2n}A_n^2, \tag{4.36}$$

where $A_n = D^2_{(x,y)}\varphi_n(t_n,x_n,y_n)$. From the viscosity supersolution property of $\tilde{W}$ to (4.29), we have

$$\begin{aligned}&\rho\tilde{W}(t_n,y_n) - p^n_{\tilde{W}} - \langle b(y_n), D_y\varphi(t_n,x_n,y_n)\rangle - \frac{1}{2}\mathrm{tr}(\sigma(y_n)\sigma^{\intercal}(y_n)N_n) \\ &\quad -\tilde{f}(t_n,y_n,\tilde{W}(t_n,y_n), -\sigma^{\intercal}(y_n)D_y\varphi(t_n,x_n,y_n)) \geq 0.\end{aligned}$$

On the other hand, from (4.34) and the viscosity subsolution property of $\tilde{V}$ to (4.26), we have

$$\begin{aligned}&\rho\tilde{V}(t_n,x_n) - p^n_{\tilde{V}} + \langle b(x_n), D_x\varphi(t_n,x_n,y_n)\rangle - \frac{1}{2}\mathrm{tr}(\sigma(x_n)\sigma^{\intercal}(x_n)M_n) \\ &\quad -\tilde{f}(t_n,x_n,\tilde{V}(t_n,x_n), \sigma^{\intercal}(x_n)D_x\varphi(t_n,x_n,y_n)) \leq 0.\end{aligned}$$

By subtracting the two previous inequalities, we obtain

$$\begin{aligned}\rho(\tilde{V}(t_n,x_n)-\tilde{W}(t_n,y_n)) \;\; \leq \;\; & p^n_{\tilde{V}} - p^n_{\tilde{W}} + \Delta F_n \\ & - \langle b(x_n), D_x\varphi_n(t_n,x_n,y_n)\rangle + \langle b(y_n), D_y\varphi_n(t_n,x_n,y_n)\rangle \\ & + \frac{1}{2}\mathrm{tr}\left(\sigma(x_n)\sigma^\intercal(x_n)M_n - \sigma(y_n)\sigma^\intercal(y_n)N_n\right), \qquad (4.37)\end{aligned}$$

where

$$\begin{aligned}\Delta F_n \;\; = \;\; & \tilde{f}(t_n,x_n,\tilde{V}(t_n,x_n),\sigma^\intercal(x_n)D_x\varphi_n(t_n,x_n,y_n)) \\ & - \tilde{f}(t_n,y_n,\tilde{W}(t_n,y_n),-\sigma^\intercal(y_n)D_y\varphi_n(t_n,x_n,y_n)).\end{aligned}$$

From (4.31), we have $p^n_{\tilde{V}} - p^n_{\tilde{W}} \to 0$ as n goes to infinity. From the Lipschitz property of b, and (4.32), we have

$$\lim_{n\to\infty}\Big(\langle b(x_n), D_x\varphi_n(t_n,x_n,y_n)\rangle + \langle b(y_n), D_y\varphi_n(t_n,x_n,y_n)\rangle\Big) \;\; = \;\; 0.$$

As usual, from (4.36), (4.31), (4.32), and the Lipschitz property of σ, we have

$$\limsup_{n\to\infty}\mathrm{tr}\left(\sigma(x_n)\sigma^\intercal(x_n)M_n - \sigma(y_n)\sigma^\intercal(y_n)N_n\right) \;\; \leq \;\; 0.$$

Moreover, by the same arguments as for $\tilde{c}$, using the nonincreasing property of $\tilde{f}$ in its third variable, and the Lipschitz property of $\tilde{f}$, we have

$$\limsup_{n\to\infty}\Delta F_n \;\; \leq \;\; 0.$$

Therefre, by sending $n \to \infty$ into (4.37), we conclude with (4.33) that $\rho(\tilde{V}-\tilde{W})(t_0,x_0) \leq 0$, a contradiction with (4.28).

• *Uniqueness for v.* The uniqueness result is then a direct consequence of the comparison principle, and the continuity of v on $[0,T)\times\mathbb{R}^d$ follows from the fact that in this case $v_* = v^*$. □

Remark 4.3 As a byproduct of the comparison principle in Theorem 4.3, we get the continuity of the value function v on $[0,T)\times\mathbb{R}^d$. Since the jump-diffusion process X is quasi-left continuous, then so is the minimal solution $Y_t = v(t,X_t)$ to the BSDE with constrained jumps, and the penalized approximation $Y^n_t = v_n(t,X_t)$. This implies that the predictable projections ${}^pY$ and ${}^pY^n$, respectively of Y and Y^n, are equal to ${}^pY_t = Y_{t^-}$ and ${}^pY^n_t = Y^n_{t^-}$. Therefore, $Y_{t^-} = \lim_{n\to\infty} Y^n_{t^-}$. From the weak version of Dini's theorem, see [9] p. 202, this yields the uniform convergence of Y^n on $[0,T]$, i.e. $\lim_{n\to\infty}\sup_{t\in[0,T]}|Y^n_t - Y_t|$ $= 0$, and so by the dominated convergence theorem, the convergence of Y^n to Y in $\mathcal{S}^\mathbf{2}$:

$$\lim_{n\to\infty}\|Y^n - Y\|_{\mathcal{S}^\mathbf{2}} \;\; = \;\; 0. \qquad (4.38)$$

Then, by applying Itô's formula to $|Y^{n+p}_t - Y^n_t|^2$, and by using similar arguments as in Lemma 3.3, one can show that

$$\|Z^{n+p}-Z^n\|^2_{\mathbf{L^2(W)}} + \|U^{n+p}-U^n\|^2_{\mathbf{L^2}(\tilde{\mu})} + \|K^{n+p}-K^n\|^2_{\mathcal{S}^\mathbf{2}} \;\; \leq \;\; C\|Y^{n+p}-Y^n\|^2_{\mathcal{S}^\mathbf{2}}.$$

Together with (4.38), this implies that $(Z^n)_n$ $(U^n)_n$ and $(K^n)_n$ are Cauchy sequences respectively in the Banach spaces $\mathbf{L^2(W)}$, $\mathbf{L^2}(\tilde{\mu})$ and $\mathcal{S}^\mathbf{2}$. Therefore, under the additional conditions **(H3)** and **(H4)** with respect to Theorem 3.1, we obtain the strong convergence of (Z^n, U^n) in $\mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu})$. Notice also that in this case, the limiting process K of K^n in $\mathcal{S}^\mathbf{2}$ is not only predictable but inherits the continuity property of K^n, see also Remark 3.3.

5 Numerical issues

In this section, we formally discuss the numerical implications of our representation results and approximation by penalization for QVIs (4.1). We first briefly recall the classical approach for numerically solving the QVI (1.1) arising from the impulse control problem (1.4). This is based on an approximation by iterated free boundary or optimal stopping problems : Starting from the function

$$u_0(t,x) \quad = \quad \mathbf{E}[g(X_T^{0,t,x}) + \int_t^T f(X_s^{0,t,x})],$$

solution to the Cauchy problem :

$$-\frac{\partial u_0}{\partial t} - \mathcal{L}u_0 - f \quad = \quad 0, \quad \text{on } [0,T) \times \mathbb{R}^d, \quad u_0(T,.) \;=\; g \quad \text{on } \mathbb{R}^d,$$

we construct the sequence of functions (u_n) by induction as the solution to the optimal stopping problem :

$$u_{n+1}(t,x) \quad = \quad \sup_{\tau \in \mathcal{T}_{t,T}} \mathbf{E}[\mathcal{H}u_n(\tau, X_\tau^{0,t,x})],$$

which satisfies the obstacle PDE :

$$\min\Big[- \frac{\partial u_{n+1}}{\partial t} - \mathcal{L}u_{n+1} - f \;,\; u_{n+1} - \mathcal{H}u_n \Big] \quad = \quad 0, \text{ on } [0,T) \times \mathbb{R}^d, \;\; u_{n+1}(T,.) \;=\; \mathcal{H}u_n \text{ on } \mathbb{R}^d,$$

where $\mathcal{H}$ is the nonlocal operator defined in (1.3). We refer to the book [16] for a more detailed description of this approximation scheme. Such an numerical approach is computationally demanding, since it requires at each induction step n, the resolution of an optimal stopping problem. Moreover, at step $n+1$, for determining the function u_{n+1} at one point, one needs to compute the function u_n in the whole space due to the nonlocal term in the obstacle $\mathcal{H}u_n$.

We consider the general QVI (4.1) and we propose here a numerical approach based on the probabilistic representation of the solution to this QVI by the constrained BSDE (2.10). We only describe the steps of the algorithm and postpone the analysis of the convergence to a future research.

Step 1. Approximation by penalized BSDE. That is, we use (3.1) to approximate (2.10). The convergence of (Y^n, Z^n, U^n, K^n) to (Y, Z, U, K) is due to Lemma 3.4 and

Theorem 3.1. We note that, by denoting $V^n_s(e) := U^n_s(e) - c(X_{s-}, Y^n_{s-}, Z^n_{s-}, e)$, (Y^n, Z^n, V^n) satisfies the following BSDE:

$$\begin{aligned} Y^n_t &= g(X_T) + \int_t^T \int_E f_n(X_s, Y^n_s, Z^n_s, V^n_s(e), e)\lambda(de)ds \\ &\quad - \int_t^T Z^n_s dW_s - \int_t^T \int_E V^n_s(e)\tilde{\mu}(de, ds), \end{aligned} \tag{5.1}$$

where

$$f_n(x, y, z, v, e) := \frac{1}{\lambda(E)} f(x, y, z) - v + nh^-(v + c(x, y, z, e), e). \tag{5.2}$$

Step 2. Discretization in E. For each m, let $E^m_1, \cdots, E^m_m$ be a partition of E the state space of the jump size. Denote $e^m_j := \frac{1}{\lambda(E^m_j)} \int_{E^m_j} e\lambda(de)$. Let $(X^m, Y^{n,m}, Z^{n,m}, V^{n,m})$ denote the solution to the following BSDE:

$$\begin{aligned} X^m_t &= x + \int_0^t b(X^m_s)ds + \int_0^t \sigma(X^m_s)dW_s + \sum_{j=1}^m \int_0^t \gamma(X^m_{s-}, e^m_j)\mu(ds, E^m_j); \\ Y^{n,m}_t &= g(X_T) + \sum_{j=1}^m \lambda(E^m_j) \int_t^T f_n(X_s, Y^{n,m}_s, Z^{n,m}_s, \Gamma^{n,m}_s(j), e^m_j)ds \\ &\quad - \int_t^T Z^{n,m}_s dW_s - \int_t^T \int_E V^{n,m}_s(e)\tilde{\mu}(ds, de), \end{aligned} \tag{5.3}$$

where

$$\Gamma^{n,m}_s(j) := \frac{1}{\lambda(E^m_j)} \int_{E^m_j} V^{n,m}_s(e)\lambda(de).$$

By the Lipschitz conditions, one can easily see that (5.3) is well-posed. Moreover, one can show that once $\max_{1\le j\le m} \mathrm{diam}(E^m_j) \to 0$ where $\mathrm{diam}(E^m_j) := \sup\{|e_1 - e_2| : e_1, e_2 \in E^m_j\}$, then

$$\lim_{m\to\infty} \Big[\|X - X^m\|_{\mathcal{S}^2} + \|Y^n - Y^{n,m}\|_{\mathcal{S}^2} + \|Z^n - Z^{n,m}\|_{\mathbf{L}^2(\mathbf{W})} + \|V^n - V^{n,m}\|_{\mathbf{L}^2(\tilde{\mu})}\Big] = 0.$$

Step 3. Discretization in time. This is an extension of the work by Bouchard and Elie [5], which studies the case $m = 1$. For any N, let $h := \frac{T}{N}$ and $t_i := ih, i = 0, \cdots, N$. First, define $X^{m,N}_{t_0} := x$, and for $i = 0, \cdots, N-1$,

$$X^{m,N}_{t_{i+1}} := X^{m,N}_{t_i} + b(X^{m,N}_{t_i})h + \sigma(X^{m,N}_{t_i})[W_{t_{i+1}} - W_{t_i}] + \sum_{j=1}^m \gamma(X^{m,N}_{t_i}, e^m_j)\mu((t_i, t_{i+1}] \times E^m_j).$$

Next, define $Y^{n,m,N}_{t_N} := g(X^{m,N}_{t_N})$, and for $i = N-1, \cdots, 0$,

$$Z^{n,m,N}_{t_i} := \frac{1}{h}\mathbf{E}_{t_i}\Big[Y^{n,m,N}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})\Big]; \tag{5.4}$$

$$\Gamma^{n,m,N}_{t_i}(j) := \frac{1}{h\lambda(E^m_j)}\mathbf{E}_{t_i}\Big[Y^{n,m,N}_{t_{i+1}}\tilde{\mu}((t_i, t_{i+1}] \times E^m_j)\Big], j = 1, \cdots, m; \tag{5.5}$$

$$Y^{n,m,N}_{t_i} = \mathbf{E}_{t_i}[Y^{n,m,N}_{t_{i+1}}] + h\sum_{j=1}^m \lambda(E^m_j) f_n(X^{m,N}_{t_i}, Y^{n,m,N}_{t_i}, Z^{n,m,N}_{t_i}, \Gamma^{n,m,N}_{t_i}(j), e^m_j) \tag{5.6}$$

Here $\mathbf{E}_{t_i}$ denotes the conditional expectation under $\mathcal{F}_{t_i}$. Notice that $Y^{n,m,N}_{t_i}$ is defined via implicit scheme. One can also define it via explicit scheme by replacing the $Y^{n,m,N}_{t_i}$ inside f_n with $Y^{n,m,N}_{t_{i+1}}$. Following the arguments in [5], one can prove the convergence of this time-discretization approximation :

$$\lim_{N\to\infty}\Big\{\max_{0\le i\le N-1}\mathbf{E}\Big[\sup_{t_i\le t\le t_{i+1}}|Y^{n,m}_t-Y^{n,m,N}_{t_i}|^2\Big]$$
$$+\sum_{i=0}^{N-1}\mathbf{E}\Big[\int_{t_i}^{t_{i+1}}[|Z^{n,m}_t-Z^{n,m,N}_{t_i}|^2+\sum_{j=1}^{m}\lambda(E^m_j)|\Gamma^{n,m}_t(j)-\Gamma^{n,m,N}_{t_i}(j)|^2]dt\Big]\Big\}=0.$$

Step 4. Approximation of the conditional expectations. The last step consists in the approximation of the conditional expectations arising in (5.4)-(5.5)-(5.6). There are several approaches proposed in the literature. We adopt here the Longstaff-Schwarz [13] method by least square projection and Monte Carlo simulations. By induction one can easily see that

$$Y^{n,m,N}_{t_i}=v^{n,m,N}_i(X^{m,N}_{t_i}),\,Z^{n,m,N}_{t_i}=\phi^{n,m,N}_i(X^{m,N}_{t_i}),\,\Gamma^{n,m,N}_{t_i}=\psi^{n,m,N}_i(X^{m,N}_{t_i}),$$

for some deterministic functions $v^{n,m,N}_i,\phi^{n,m,N}_i,\psi^{n,m,N}_i$. For any L, choose some basis functions $(v^L_l,\phi^L_l,\psi^L_l), l=1,\cdots,L$. For any M, simulate M independent copies of $W^k_{t_{i+1}}-W^k_{t_i}$ and $\tilde{\mu}^k((t_i,t_{i+1}]\times E^m_j)$, $k=1,\cdots,M$.

First, for each k, define $X^{m,N,M}_{k,t_0}:=x$, and for $i=0,\cdots,N-1$,

$$\begin{aligned}X^{m,N,M}_{k,t_{i+1}} &:= X^{m,N,M}_{k,t_i}+b(X^{m,N}_{k,t_i})h+\sigma(X^{m,N,M}_{k,t_i})[W^k_{t_{i+1}}-W^k_{t_i}]\\ &\quad+\sum_{j=1}^{m}\gamma(X^{m,N,M}_{k,t_i},e^m_j)\mu((t_i,t_{i+1}]\times E^m_j).\end{aligned}$$

Next, define $Y^{n,m,N,L,M}_{k,t_N}:=g(X^{m,N,M}_{k,t_N})$ for each k, and for $i=N-1,\cdots,0$, we define $(Z^{n,m,N,L,M}_{k,t_i},\Gamma^{n,m,N,L,M}_{k,t_i},Y^{n,m,N,L,M}_{k,t_i})$ as follows.

$$(\hat{\alpha}_1,\cdots,\hat{\alpha}_L):=\arg\min_{\alpha_1,\cdots,\alpha_L}\frac{1}{M}\sum_{k=1}^{M}\Big|\frac{1}{h}Y^{n,m,N,L,M}_{k,t_{i+1}}[W^k_{t_{i+1}}-W^k_{t_i}]-\sum_{l=1}^{L}\alpha_l\phi^L_l(X^{m,N,M}_{k,t_i})\Big|^2;$$

$$Z^{n,m,N,L,M}_{k,t_i}:=\sum_{l=1}^{L}\hat{\alpha}_l\phi^L_l(X^{m,N,M}_{k,t_i});$$

$$(\hat{\beta}_1(j),\cdots,\hat{\beta}_L(j)):=\arg\min_{\beta_1,\cdots,\beta_L}\frac{1}{M}\times$$
$$\sum_{k=1}^{M}\Big|\frac{1}{h\lambda(E^m_j)}Y^{n,m,N,L,M}_{t_{i+1}}\tilde{\mu}^k((t_i,t_{i+1}]\times E^m_j)-\sum_{l=1}^{L}\beta_l\psi^L_l(X^{m,N,M}_{k,t_i})\Big|^2;$$

$$\Gamma^{n,m,N,L,M}_{k,t_i}(j):=\sum_{l=1}^{L}\hat{\beta}_l(j)\psi^L_l(X^{m,N,M}_{k,t_i});$$

$$(\hat{\gamma}_1,\cdots,\hat{\gamma}_L):=\arg\min_{\gamma_1,\cdots,\gamma_L}\frac{1}{M}\sum_{k=1}^{M}\Big|Y^{n,m,N,L,M}_{k,t_{i+1}}-\sum_{l=1}^{L}\gamma_l v^L_l(X^{m,N,M}_{k,t_i})\Big|^2;$$

$$\tilde{Y}^{n,m,N,L,M}_{k,t_i} := \sum_{l=1}^{L} \hat{\gamma}_l v^L_l(X^{m,N,M}_{k,t_i});$$

Finally we solve the following equation for $Y^{n,m,N,L,M}_{k,t_i}$ via Picard iteration:

$$Y^{n,m,N,L,M}_{k,t_i} = \tilde{Y}^{n,m,N,L,M}_{k,t_i} + h\sum_{j=1}^{m} \lambda(E^m_j) f_n(X^{m,N,M}_{k,t_i}, Y^{n,m,N,L,M}_{k,t_i}, Z^{n,m,N,L,M}_{k,t_i}, \Gamma^{n,m,N,L,M}_{k,t_i}(j), e^m_j).$$

6 Some sufficient conditions for (H2') and (H3)

In this section, we provide various explicit conditions on the coefficients model, which ensure that the general assumptions **(H2')** and **(H3)** hold true.

6.1 Existence of the solution to BSDE with jump constraint

We first consider a case where we have upper bounds for the coefficients and $h(u,e) = -u$.

Proposition 6.1 *Suppose that $h(u,e) = -u$, and assume that there exist real constants C_1, C_2 and $\eta \in \mathbb{R}^d$ such that*

$$g(x) \le C_1 + \langle \eta, x\rangle \ , \ c(x,y,z,e) + \langle \eta, \gamma(x)\rangle \le 0 \quad \textit{and} \quad f(x,y,z) + \langle \eta, b(x)\rangle \le C_2, \tag{6.7}$$

for all $(x,y,z,e) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times E$. Then **(H2')** *holds true.*

Proof. Let us define a quadruple $(\tilde{Y}, \tilde{Z}, \tilde{K}, \tilde{U})$ by : $\tilde{Y}_t = C_1 + C_2(T-t) + \langle \eta, X_t\rangle$ for $t < T$, $\tilde{Y}_T = g(X_T)$, $\tilde{Z}_t = \sigma(X_{t^-}).\eta$, $\tilde{U}_t(e) = 0$ and

$$\begin{aligned}
\tilde{K}_t &= \int_0^t \Big\{C_2 - \eta\cdot b(X_s) - f(X_s, \tilde{Y}_s, \tilde{Z}_s)\Big\} ds \\
&\quad - \int_0^t \int_E \Big\{c(X_{s-}, \tilde{Y}_{s-}, \tilde{Z}_s, e) + \langle \eta, \gamma(X_{s-})\rangle\Big\} \mu(ds,de), \quad t < T, \\
\tilde{K}_T &= \tilde{K}_{T-} + C_1 + \langle \eta, X_T\rangle - g(X_T).
\end{aligned}$$

From (6.7), the process $\tilde{K}$ is clearly nondecreasing. Moreover, from the dynamics of X, and by construction, we see that the quadruple $(\tilde{Y}, \tilde{Z}, \tilde{K}, \tilde{U})$ satisfies (2.10)-(2.13) and the function $\tilde{v}(t,x) = C_1 + C_2(T-t) + \eta.x$ clearly satisfies a linear growth condition. □

We next give an example inspired by [4] where the jumps of X vanish as X goes out of a ball centered in zero in the case of impulse control.

Proposition 6.2 *Suppose that $h(u,e) = -u$, f, c does not depend on y, z, and assume that $c \le 0$, $\gamma = 0$ on $\{x \in \mathbb{R}^d \ : \ |x| \ge C_1\} \times E$ for some $C_1 > 0$. Then,* **(H2')** *holds true.*

Proof. We consider the function v :

$$v(t,x) = \sup_{\nu\in\mathcal{V}} \mathbf{E}^\nu\Big[g(X^{t,x}_T) + \int_t^T f(X^{t,x}_s) ds + \int_t^T \int_E c(X^{t,x}_{s^-}, e)\mu(ds,de)\Big].$$

Since $c \leq 0$, and the choice of $\nu = 1$ corresponds to the probability measure $\mathbf{P}^1 = \mathbf{P}$, we see that $\hat{v} \leq v \leq \bar{v}$ where

$$
\begin{aligned}
\hat{v}(t,x) &= \mathbf{E}\Big[g(X_T^{t,x}) + \int_t^T f(X_s^{t,x})ds + \int_t^T \int_E c(X_{s^-}^{t,x}, e)\mu(ds,de)\Big] \\
\bar{v}(t,x) &= \sup_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T^{t,x}) + \int_t^T f(X_s^{t,x})ds\Big].
\end{aligned}
$$

The function $\hat{v}$ clearly satisfies a linear growth condition by the linear growth conditions on g, f, c and the standard estimate for X. Moreover, under the assumptions on the jump coefficient γ, it is shown in [4] that $\bar{v}$ satisfies a linear growth condition. Therefore, $\hat{v}$ also satisfies a linear growth condition.

Let us now define the process $Y_t = v(t, X_t)$, which is then equal to

$$
Y_t = \operatorname*{ess\,sup}_{\nu \in \mathcal{V}} \mathbf{E}^\nu \Big[g(X_T) + \int_t^T f(X_s)ds + \int_t^T \int_E c(X_{s^-}, e)\mu(ds,de)\Big|\mathcal{F}_t\Big],
$$

and lies in $\mathcal{S}^\mathbf{2}$ from the linear growth condition, ad the estimate (2.2) for X. From Theorem 2.1, we then know that there exists $(Z,U,K) \in \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$ such that (Y,Z,U,K) is the minimal solution to (2.10)-(2.13), and so **(H2')** is satisfied. □

We finally consider a case for general constraint function h.

Proposition 6.3 *Assume that there exists a Lipschitz function $w \in \mathcal{C}^2(\mathbb{R}^d)$ satisfying a linear growth condition, supersolution to (4.3), and such that*

$$
\langle b, Dw \rangle + \frac{1}{2} tr(\sigma\sigma^\intercal D^2 w) + f(\cdot, w, \sigma^\intercal Dw) \quad \leq \quad C, \qquad on \ \mathbb{R}^d,
$$

for some constant C. Then **(H2')** *holds true.*

Proof. Let us define a quadruple $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K})$ by

$$
\tilde{Y}_t = w(X_t) + C(T-t), \quad t < T, \qquad \tilde{Y}_T = g(X_T),
$$

$\tilde{Z}_t = \sigma^\intercal(X_{t^-})Dw(X_{t^-})$, $\tilde{U}_t(e) = w(X_{t^-} + \gamma(X_{t^-}, e)) + c(X_{t^-}, \tilde{Y}_{t^-}, \tilde{Z}_t, e) - w(X_{t^-})$, and

$$
\begin{aligned}
\tilde{K}_t &= \int_0^t [C - \langle b(X_s), Dw(X_s)\rangle - \frac{1}{2}\mathrm{tr}\{\sigma(X_s)\sigma^\intercal(X_s)D^2 w(X_s)\} - f(X_s, \tilde{Y}_s, \tilde{Z}_s)]ds, \quad t < T, \\
\tilde{K}_T &= \tilde{K}_{T^-} + w(X_T) - g(X_T).
\end{aligned}
$$

From the conditions on w, we see that $(\tilde{Y}, \tilde{Z}, \tilde{K}, \tilde{U})$ lies in $\mathcal{S}^\mathbf{2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{A^2}$. Moreover, by Itô's formula to $w(X_t)$ and the supersolution property of w to (4.3), we conclude that $(\tilde{Y}, \tilde{Z}, \tilde{K}, \tilde{U})$ is solution to (2.10)-(2.11), and $\tilde{v}(t,x) = w(t,x) + C(T-t)$ satisfies a linear growth condition. □

6.2 The strict supersolution condition (H3)

We give a sufficient condition for **(H3)** in the usual case where f and c do not depend neither on y nor on z.

Proposition 6.4 *Consider the case where h is given by*

$$h(u, e) = -u.$$

Assume that there exists a constant $\alpha > 0$ such that

$$\begin{aligned} -\alpha &< |x + \gamma(x, e)|^2 - |x|^2 \quad \forall (x, e) \in \mathbb{R}^d \times E \\ \beta &:= \inf_{(x,e) \in \mathbb{R}^d \times E} \frac{-c(x, e)}{|x + \gamma(x, e)|^2 - |x|^2 + \alpha} > 0 \end{aligned}$$

Then assumption **(H3)** *holds true.*

Proof. We set $\Lambda(x) := \beta|x|^2 + \zeta$ with ζ large enough so that $\Lambda \geq g$, i.e. **(H3)**(iii) is satisfied. A straightforward computation shows that

$$\inf_{e \in E} h(\mathcal{H}^e \Lambda(x) - \Lambda(x), e) \geq \alpha\beta > 0$$

and hence **(H3)** (ii) is satisfied. Clearly, **(H3)** (iv) holds as well. Finally, it follows from the linear growth assumption on b and σ that **(H3)** (i) holds for a sufficiently large parameter ρ. □

References

[1] Barles G. (1994) : *Solutions de viscosité des équations d'Hamilton-Jacobi*, Mathématiques et Applications, Springer Verlag.

[2] Barles G., Buckdahn R. and E. Pardoux (1997) : "Backward stochastic differential equations and integral-partial differential equations", *Stochastics and Stochastics Reports*, **60**, 57-83.

[3] Bensoussan A. and J.L. Lions (1984) : *Impulse control and quasi-variational inequalities*, Gauthier-Villars.

[4] Bouchard B. (2006) : "A stochastic target formulation for optimal switching problems in finite horizon", Preprint.

[5] Bouchard B. and R. Elie (2008) : "Discrete-time approximation of decoupled forward-backward SDE with jumps", *Stochastic Processes and their Applications*, **118**, 53-75.

[6] Buckdahn R. and Y. Hu (1998) : "Hedging contingent claims for a large investor in an incomplete market", *Advances in Applied Probability*, **30**, 239-255.

[7] Crandall M., Ishii H. and P.L. Lions (1992) : "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.*, **27**, 1-67.

[8] Cvitanic J., Karatzas I and M. Soner (1998) : "Backward stochastic differential equations with constraints on the gain-process", *Annals of Probability*, **26**, 1522-1551.

[9] Dellacherie C. and P.A. Meyer (1980) : *Probabilités et Potentiel, V-VIII*, Hermann, Paris.

[10] El Karoui N., Kapoudjian C., Pardoux E., Peng S. and M.C. Quenez (1997) : "Reflected solutions of Backward SDE's, and related obstacle problems for PDEs", *Annals of Probability*, **25**, 702-737.

[11] El Karoui N. and M.C. Quenez (1995) : "Dynamic programming and pricing of contingent claims in an incomplete market", *SIAM J. Control and Optimization*, **33**, 29-66.

[12] Hamadène S. and Y. Ouknine (2003) : "Reflected backward stochastic differential equations with jumps and random obstacle", *Elect. Journal of Prob.*, **8**, 1-20.

[13] Longstaff F. and E. Schwartz (1998) : "Valuing American options by simulation : a least-squares approach", *Review of Financial Studies*, **14**, 113-147.

[14] Ma, J. and J. Zhang (2002) : "Path Regularity for Solutions of Backward SDE's", *Proba. Theory and Relat. Fields*, **122**, 163-190.

[15] Meyer, P. and W. Zheng (1984) : "Tightness criteria for laws of semimartingales", *Ann. Inst. Henri Poincaré*, **20**, 353-372.

[16] Oksendal B. and A. Sulem (2007) : *Applied stochastic control of jump-diffusions*, Universitext, Springer Verlag, 2nd edition.

[17] Pardoux E. and S. Peng (1990) : "Adapted solution of a backward stochastic differential equation", *Systems and Control Letters*, **14**, 55-61.

[18] Peng S. (1999) : "Monotonic limit theory of BSDE and nonlinear decomposition theorem of Doob-Meyer's type", *Prob. Theory and Related Fields*, **113**, 473-499.

[19] Peng S. and M. Xu (2007) : "Constrained BSDE and viscosity solutions of variation inequalities", Preprint.

[20] Royer M. (2006) : "Backward stochastic differential equations with jumps and related nonlinear expectations", *Stochastic Processes and their Applications*, **116**, 1358-1376.

[21] Tang S. and X. Li (1994) : "Necessary conditions for optimal control of stochastic systems with jumps", *SIAM J. Control and Optimization*, **32**, 1447-1475.